

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2004, 14:42:06 ; Search time 79 Seconds

(without alignments)
2929.409 Million cell updates/sec

Title: US-10-054-691-2

Perfect score: 7766
Sequence: 1 MGLRPGIFLELLILLIGQT.....RCRRGRREDPPMSLRVAL 1458

Scoring table: BLOSUM62

Gapop 10.0, Gapept 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: A_Geneseq_19Jun03.*

```
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	7766	100.0	1458	23	ABBO9555 Human lipase NHL (
2	7762	99.9	1458	23	ABBO9556 Human lipase NHL (
3	7759	99.9	1458	24	AAE34440 Human lipid-associ
4	7383	95.1	1433	23	ABU65083 Human NOV24c prote
5	6311	81.3	1419	23	ABU65081 Human NOV24a prote
6	6253	80.5	1216	23	AAE22860 Human phospholipas
7	5279	68.0	1450	18	AAW20751 Rat phospholipase-
8	4911	63.2	969	24	AAE34448 Human lipid-associ
9	2492	32.1	472	23	ABP53556 Human phospholipas

10	1525	19.6	310	23	ABU65082 Human NOV24b prote
11	1360	17.5	267	22	ABBI1053 Human phospholipas
12	1360	17.5	267	22	AAW25824 Human protein sequ
13	584	7.5	424	22	ABBI1556 Drosophila melanog
14	560	7.2	148	22	ABBI6111 Human testicular a
15	560	7.2	148	22	AAW95420 Human reproductive
16	527	6.8	101	23	ABBI9352 Human polypeptide
17	515	6.6	447	22	ABBI5406 Drosophila melanog
18	459	5.9	132	22	ABBI1237 Human phospholipas
19	324	4.2	334	22	ABGI2997 Novel human diagno
20	249	3.2	109	22	AAE10214 Human bone marrow
21	178	2.3	102	23	ABP34826 Human ORF3799 prot
22	138	1.8	10182	23	ABP38314 Staphylococcus epi
23	137	1.8	639	22	ABBI6805 Drosophila melanog
24	137	1.8	1369	22	ABBI58070 Drosophila melanog
25	136	1.8	1424	23	ABGI6726 Human novel polype
26	136	1.8	1455	19	AAW48663 Fancconi anaemia of
27	132	1.7	2542	22	ABBI1137 Drosophila melanog
28	129	1.7	1218	22	AAE62801 DNA ligase II amin
29	128	1.6	589	23	ABBI3511 Lactococcus lactis
30	127	1.6	1016	19	AAW41312 CF-5 pathogen resl
31	127	1.6	787	24	ABJ26529 Aspergillus fumiga
32	125	1.6	1339	23	ABBO6077 Human NS protein s
33	125	1.6	30	18	AAW30753 Phospholipase-B/11
34	124	1.6	1178	18	AAW30763 Mannose-1-phosphat
35	124	1.6	2828	23	ABGI2897 Human Adlican prot
36	124	1.6	2828	23	ABGI3139 Human Adlican prot
37	124	1.6	2828	23	ABGI3132 Human Adlican-2 pr
38	124	1.6	2828	24	ABR7439 Human breast asso
39	124	1.6	2828	24	ABJ37051 Human breast cance
40	124	1.6	2828	24	ABUS6609 Lung cancer-associ
41	123	1.6	1230	19	AAW98275 H. pylori GHP0 690
42	122	1.6	744	22	ABBI1226 Drosophila melanog
43	121	1.6	1031	22	ABBI5127 Drosophila melanog
44	121	1.6	1097	22	ABBI1579 Drosophila melanog
45	121	1.6	2597	21	AAV53664 Mechanical stress

ALIGNMENTS

RESULT 1	ABB09555 standard; Protein; 1458 AA.
ID	ABB09555
XX	ABB09555;
AC	21-OCT-2002 (first entry)
XX	
DE	Human Lipase NHL (Ala 1318 variant).
XX	
KW	Human, lipase; chromosome 2; phospholipase B homologue; thyroid; brain;
KW	inflammatory disease; proliferative disease; infectious disease;
KW	clotting disorder; cancer; drug screening; mental disorder; NHL;
KW	diagnostic reagent; clinical trial monitoring; cosmetic; nutraceutical;
KW	mutation detection; gene expression analysis; transgenic animal;
KW	neurotropic; cytostatic; antiinflammatory; single nucleotide polymorphism;
KW	SNP; gene therapy; enzyme.
XX	
OS	Homo sapiens.
XX	
FT	Key
FT	Misc-difference 1318
FT	Location/Qualifiers
FT	/note="Ala replaces Ala in a polymorphic variant (ABB09555)"
PN	WO200259328-A1.
XX	
PD	01-AUG-2002.
XX	
PF	22-JAN-2002; 2002WO-US01715.
XX	
PR	24-JAN-2001; 2001US-264049P.

XX (LEXI-) LEXICON GENETICS INC.
 PA Yu X, Miranda M, Turner CA;
 PI WPI; 2002-599797/64.
 DR N-PSDB; ABQ77623.
 XX
 PT Polynucleotides encoding human lipases that are structurally related to
 PT animal lipases, particularly phospholipase B, useful for drug
 PT screening, diagnosis and in gene therapy of biological disorders -
 PS Claim 2; Page 37-41; 44pp; English.
 XX
 CC The invention relates to a novel human lipase (NHL; ABB09555, ABB09556)
 CC and to nucleic acids encoding it (ABQ77623, ABQ77624). The NHL has
 CC structural similarity with animal lipases, particularly phospholipase B.
 CC Polynucleotides encoding NHL were obtained using human genomic sequences
 CC in conjunction with human thyroid and brain cDNAs. The NHL gene is
 CC located on chromosome 2, and contains a C/T polymorphism at position
 CC 3953 of the open reading frame (ORF), resulting in an Ala/Val
 CC substitution at position 1318 in the protein. NHL nucleotides and
 CC proteins are useful for treating disorders such as inflammatory or
 CC proliferative disease, infectious disease, clotting disorders, and
 CC cancer. They can also be used in screening for compounds useful in
 CC the treatment of mental, biological or medical disorders, as diagnostic
 CC reagents, in clinical trial monitoring and in cosmetic and nutritional
 CC applications. NHL nucleotides can additionally be used in the detection
 CC of disease-associated mutations, in the analysis of gene expression, for
 CC the recombinant expression of NHL, to generate transgenic animals, in
 CC gene therapy, and as part of ribozyme and/or triple helix sequences
 CC useful in the modulation of NHL gene expression. The present sequence
 CC represents the Ala 1318 variant of NHL.

XX Sequence 1458 AA;

Query Match 100.0%; Score 7766; DB 23; Length 1458;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLRPGIFLELLLELLGQGTPOIHTSPKSTLSEGLMPELTKNSPPPCNPNKLGVMMSK 60
 DB 1 MGLRPGIFLELLLELLGQGTPOIHTSPKSTLSEGLMPELTKNSPPPCNPNKLGVMMSK 60
 QY 61 SVSLKRSQDIFKFAAIGNLEIPDPGTCGLDEKQDWTERRPOQVCGVNTVSDIIRYSPS 120
 DB 61 SVSLKRSQDIFKFAAIGNLEIPDPGTCGLDEKQDWTERRPOQVCGVNTVSDIIRYSPS 120
 QY 121 VPMPCVCHTGRVTPHDAEDLMTIOAELVNMKNENLQDFOPMKLINVFESNASQCYLC 180
 DB 121 VPMPCVCHTGRVTPHDAEDLMTIOAELVNMKNENLQDFOPMKLINVFESNASQCYLC 180
 QY 181 PSAQNGLAAGVDELMLGVLDYLQOEVPRAFVNILVDLSEVAEVSROHYGTWLSAPAPPCN 240
 DB 181 PSAQNGLAAGVDELMLGVLDYLQOEVPRAFVNILVDLSEVAEVSROHYGTWLSAPAPPCN 240
 QY 241 CSEETTRILAVNMQMSQEAHNSGLASSRSYSEQSPFVVQPFYETTPSLHSEDPLOD 300
 DB 241 CSEETTRILAVNMQMSQEAHNSGLASSRSYSEQSPFVVQPFYETTPSLHSEDPLOD 300
 QY 301 STTLAHLNMRMMPAGEKDEPLSVKHGRPMKCPQSPFLFSYRNSNYLTRLOKPODKL 360
 DB 301 STTLAHLNMRMMPAGEKDEPLSVKHGRPMKCPQSPFLFSYRNSNYLTRLOKPODKL 360
 QY 361 EVREGAIRCPDPSDTPVTSVHRLKPADINVIGALGDSLTAGNGAGSTPGANVLDTLTQ 420
 DB 361 EVREGAIRCPDPSDTPVTSVHRLKPADINVIGALGDSLTAGNGAGSTPGANVLDTLTQ 420
 QY 421 YRGLSNVSGDENIGTITTLANILREFNPSLKGFSVGTGKETSBNAPLNOAVAGRAEDL 480
 DB 421 YRGLSNVSGDENIGTITTLANILREFNPSLKGFSVGTGKETSBNAPLNOAVAGRAEDL 480
 QY 481 PVGARRLVDMKNDTRHFOEDMKITLFIGGNDLDFCNDLVHYSPQNTDNI GKALDI 540

DB 481 PVGARRLVDMKNDTRHFOEDMKITLFIGGNDLDFCNDLVHYSPQNTDNI GKALDI 540
 QY 541 LHAEPRAFVNILVTLEIVNLRELVOEKVYCPRMILRSICPCVLKFDNDSTELATLIEF 600
 DB 541 LHAEPRAFVNILVTLEIVNLRELVOEKVYCPRMILRSICPCVLKFDNDSTELATLIEF 600
 QY 601 NKKFOEKTOLISSGRDTRDEFTVVQPFENVDMPTSEGLPDNSFFPDCPFSSKS 660
 DB 601 NKKFOEKTOLISSGRDTRDEFTVVQPFENVDMPTSEGLPDNSFFPDCPFSSKS 660
 QY 661 HSRASALMNMNMLEPVQOKTTRKFKENKINITCPNOVOPFLRTYKMSQGHGTMPCDR 720
 DB 661 HSRASALMNMNMLEPVQOKTTRKFKENKINITCPNOVOPFLRTYKMSQGHGTMPCDR 720
 QY 721 APSALHPTSVHALRPADIOVVAALGSLTRNGCIGSKPDDLPTVTOYRGLSAGDGS 780
 DB 721 APSALHPTSVHALRPADIOVVAALGSLTRNGCIGSKPDDLPTVTOYRGLSAGDGS 780
 QY 781 LENVTLPLNLRREFNRLTGYAVGTGDANDTNAPLNOAVGAKAEDLMSOVOTLMQMKD 840
 DB 781 LENVTLPLNLRREFNRLTGYAVGTGDANDTNAPLNOAVGAKAEDLMSOVOTLMQMKD 840
 QY 841 DHRVNFHEDMKVITVLIGGSDLCDYCTDSNLYSAANFVDHLARNALDVLHREVPRLVNLV 900
 DB 841 DHRVNFHEDMKVITVLIGGSDLCDYCTDSNLYSAANFVDHLARNALDVLHREVPRLVNLV 900
 QY 901 DFLNPTIMROVFLGPNPDKCPVOQASVLCNVLTLRENSOELARLEAFSRAYSRSMELVG 960
 DB 901 DFLNPTIMROVFLGPNPDKCPVOQASVLCNVLTLRENSOELARLEAFSRAYSRSMELVG 960
 QY 961 SGRVYDQEDSVYLQPFQFQNICLPVLDLGPDSFAPDCITHNOKFHSGLAALMTNML 1020
 DB 961 SGRVYDQEDSVYLQPFQFQNICLPVLDLGPDSFAPDCITHNOKFHSGLAALMTNML 1020
 QY 1021 EPLGSKTETLDLRAEMPTICTPONEBFLRTPRNSNTYPIKPAIENMGSDFLCTEWRASN 1080
 DB 1021 EPLGSKTETLDLRAEMPTICTPONEBFLRTPRNSNTYPIKPAIENMGSDFLCTEWRASN 1080
 QY 1081 SVPTSVHQLRPADIKVVAALGDSLTTAVGARPNNSDDLPTSMRGLSWSIGDGNLETHTT 1140
 DB 1081 SVPTSVHQLRPADIKVVAALGDSLTTAVGARPNNSDDLPTSMRGLSWSIGDGNLETHTT 1140
 QY 1141 LPMILKKFNPYLLGFSTSTWEGTAGLVAAEGARADMPOANDLVERMKNSPDINLEKD 1200
 DB 1141 LPMILKKFNPYLLGFSTSTWEGTAGLVAAEGARADMPOANDLVERMKNSPDINLEKD 1200
 QY 1201 WKLVTLFIGVNDLCHYCENPEALATEYVOHIOQALDISEELPRAFVNVEVMEIASLY 1260
 DB 1201 WKLVTLFIGVNDLCHYCENPEALATEYVOHIOQALDISEELPRAFVNVEVMEIASLY 1260
 QY 1261 OGOGGKCAMLAANNCTCLRHSSQSLSEKOLKKNVNNLGHGSSFSYWHQYTORDEPAVV 1320
 DB 1261 OGOGGKCAMLAANNCTCLRHSSQSLSEKOLKKNVNNLGHGSSFSYWHQYTORDEPAVV 1320
 QY 1321 VQPFQNTLPLNERDPTDLPFSEDCPFHSDGHAEMALMNMNLEPGRKTTSNFT 1380
 DB 1321 VQPFQNTLPLNERDPTDLPFSEDCPFHSDGHAEMALMNMNLEPGRKTTSNFT 1380
 QY 1381 HSRAKLKCPSPSPYLYTLRNSRLPDQAEAPBEVLYMAVPVAGVGLVGIIGTVVWRC 1440
 DB 1381 HSRAKLKCPSPSPYLYTLRNSRLPDQAEAPBEVLYMAVPVAGVGLVGIIGTVVWRC 1440
 QY 1441 RRGGRREDPMSLRVAL 1458
 DB 1441 RRGGRREDPMSLRVAL 1458

RESULT 2
 ABB09556
 ID ABB09556 standard; Protein; 1458 AA.
 XX
 AC ABB09556;

QY 1141 LPNLLKFNFFLLGFTSTWGTAGLVAAEGARAPDPAQAMDLVERMKNSPDINLEKD 1200
 DB 1141 LPNLLKFNFFLLGFTSTWGTAGLVAAEGARAPDPAQAMDLVERMKNSPDINLEKD 1200
 QY 1201 WKLVTFIIGVNDLCHYCEPBAHLATEYVOHIOQALDIISEELPRAFVNVVEWELASLY 1260
 DB 1201 WKLVTFIIGVNDLCHYCEPBAHLATEYVOHIOQALDIISEELPRAFVNVVEWELASLY 1260
 QY 1261 QGGGKCAMLAANNCTCLHSSOSLEKOLKKNMMLQHGISFSYWHQYTOREDPAVY 1320
 DB 1261 QGGGKCAMLAANNCTCLHSSOSLEKOLKKNMMLQHGISFSYWHQYTOREDPAVY 1320
 QY 1321 VQPFQNTLPLNERGDTDLTFSEDCFHSRSGHAAEMALMNMMLPEYGRKTTNNFT 1380
 DB 1321 VQPFQNTLPLNERGDTDLTFSEDCFHSRSGHAAEMALMNMMLPEYGRKTTNNFT 1380
 QY 1381 HSRBAKLKCPSPSPYLYTLNRSRLPDQAEAPVLYAVVAGVGVGIICTVVMRC 1440
 DB 1381 HSRBAKLKCPSPSPYLYTLNRSRLPDQAEAPVLYAVVAGVGVGIICTVVMRC 1440
 QY 1441 RRGGRREDPPMSLRTVAL 1458
 DB 1441 RRGGRREDPPMSLRTVAL 1458

RESULT 3
 ID AAE34440 standard; Protein; 1458 AA.
 XX AAE34440;
 XX 14-MAY-2003 (first entry)
 DE Human lipid-associated molecule (LIPAM)-1 protein.
 XX Human; lipid-associated molecule; LIPAM; cardiovascular disorder; stroke;
 KW arteriovenous fistula; atherosclerosis; hypertension; Raynaud's disease;
 KW aneurysm; congestive heart failure; thrombophlebitis; angina pectoris;
 KW ischaemic heart disease; rheumatic heart disease; peptic esophagitis;
 KW gastrointestinal disorder; lipid metabolism disorder; Crohn's disease;
 KW nausea; peptic ulcer; fatty liver; Fabry's disease; Gaucher's disease;
 KW diabetes mellitus; hyperlipidaemia; hypercholesterolaemia; epilepsy;
 KW autoimmune disorder; inflammatory disorder; neurological disorder; kuru;
 KW acquired immunodeficiency syndrome; anaemia; Alzheimer's disease; asthma;
 KW dementia; prion disease; Creutzfeldt-Jakob disease; leukaemia; cancer;
 KW adenocarcinoma; lymphoma; melanoma; myeloma; sarcoma; gene therapy;
 KW protein replacement therapy.
 XX Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..19
 FT Protein /label= Signal-peptide
 FT /note= "Human mature LIPAM-1 protein"
 FT Peptide 1..21
 FT Protein /label= Signal-peptide
 FT /note= "Human mature LIPAM-1 protein"
 FT Peptide 1..23
 FT Protein /label= Signal-peptide
 FT /note= "Human mature LIPAM-1 protein"
 FT Peptide 1..27
 FT Protein /label= Signal-peptide
 FT /note= "Human mature LIPAM-1 protein"
 FT Peptide 28..1458
 FT Protein /note= "Human mature LIPAM-1 protein"
 FT 393..521
 FT Domain /note= "Lipase/acylhydrolase with GDSL-like motif"

FT Active-site 394..404
 FT /note= "Lipolytic enzyme G-D-S-L family
 FT serine active-site"
 FT Domain 740..868
 FT /note= "Lipase/acylhydrolase with GDSL-like motif"
 FT Active-site 741..751
 FT /note= "Lipolytic enzyme G-D-S-L family
 FT serine active-site"
 FT Domain 1096..1219
 FT /note= "Lipase/acylhydrolase with GDSL-like motif"
 FT Domain 1415..1442
 FT /note= "Transmembrane domain"
 PN WC0200294988-A2.
 XX
 XX 28-NOV-2002.
 PD
 XX
 XX 17-MAY-2002; 2002MO-US15688.
 PF
 XX
 PR 18-MAY-2001; 2001US-292242P.
 PR 25-MAY-2001; 2001US-293726P.
 PR 01-JUN-2001; 2001US-295346P.
 PR 06-JUL-2001; 2001US-303404P.
 PR 24-AUG-2001; 2001US-314754P.
 PR 22-JAN-2002; 2002US-351262P.
 PR 29-MAR-2002; 2002US-368799P.
 XX
 XX (INCYT-) INCYTE GENOMICS INC.
 PA
 XX Tang YT, Yue H, Azimzai Y, Baughn MR, Burford N, Reddy R,
 PI Wallia NK, Das D, Nguyen DB, Yao MG, Arvizu CS, Lu Y, Gandhi AR,
 PI Griffitt JA, Elliott VS, Ramkumar J, Lal PG, Lu DM, Lee EA,
 PI Lee SY, Yue H, Yang J, Tribouley CM, Kable AE, Swarnakar A,
 XX WPI, 2003-120797/11.
 DR N-PSDB; AAD52626.
 DR
 XX New human lipid-associated molecule (LIPAM) proteins and
 PT polynucleotides, useful for diagnosing, treating or preventing
 PT cardiovascular disorders (e.g. aneurysms), neurological disorders (e.g.
 PT Parkinson's disease) or cancers -
 XX
 PS Claim 56, Page 137-141; 171pp; English.
 XX
 CC The present invention relates to novel human lipid-associated molecules
 CC (LIPAM) and polynucleotides encoding such proteins. Sequences of the
 CC invention are useful for treating diseases or conditions associated with
 CC decreased expression of functional LIPAM. The antagonist is useful for
 CC treating a disease or condition associated with the overexpression of
 CC functional LIPAM. They are useful for diagnosing, treating or preventing
 CC cardiovascular disorders (e.g. arteriovenous fistula, atherosclerosis,
 CC hypertension, Raynaud's disease, aneurysms, varicose veins, congestive
 CC heart failure, thrombophlebitis, angina pectoris, ischaemic heart disease
 CC or rheumatic heart disease), gastrointestinal disorders (e.g. peptic
 CC esophagitis, nausea, peptic ulcer or Crohn's disease), lipid metabolism
 CC disorders (e.g. fatty liver, Fabry's disease, Gaucher's disease, diabetes
 CC mellitus, hyperlipidaemia, hypercholesterolaemia), autoimmune disorders
 CC or inflammatory disorders (e.g. acquired immunodeficiency syndrome,
 CC anaemia, asthma or Crohn's disease), neurological disorders (e.g. stroke,
 CC epilepsy, dementia, Alzheimer's disease, or prion diseases such as kuru
 CC or Creutzfeldt-Jakob disease) or cancers (e.g. adenocarcinoma, leukaemia,
 CC lymphoma, melanoma, myeloma or sarcoma). They are also used in gene
 CC therapy and protein replacement therapy. The present sequence is human
 CC LIPAM-1 protein.
 CC
 XX Sequence 1458 AA;
 XX
 Query Match 99.9%; Score 7759; DB 24; Length 1458;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1457; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MGLRPGIFLLELLLLGQGTPTQITSPRKSTLGGOLMPETLKNSPPPCNBNKLGVNMPK 60
 |||||||

Db 1 MGRLPGIPELLLELLLLLGQTPQIHTSPKSTLEBQIMPELTKNSPPECNPKLGVMNSK 60
 QY 61 SVHSLSKSDIKFVAAGNLBIIPDPGTGDELEKODMTERPOOVGVWTVLSDIIRYSPS 120
 Db 61 SVHSLSKSDIKFVAAGNLBIIPDPGTGDELEKODMTERPOOVGVWTVLSDIIRYSPS 120
 QY 121 VPMVCHTGKRVIPHDGAEDLMIOAQLVNMKNELQJDFQDMKLIINVEFSNASCYLC 180
 Db 121 VPMVCHTGKRVIPHDGAEDLMIOAQLVNMKNELQJDFQDMKLIINVEFSNASCYLC 180
 QY 181 PSAQONGLAAGVDELGMVLDYLOQELPRAFVNLVDLSVAEVSROYHGTWLSAPAPPCN 240
 Db 181 PSAQONGLAAGVDELGMVLDYLOQELPRAFVNLVDLSVAEVSROYHGTWLSAPAPPCN 240
 QY 241 CSSEETRLAVYVQWMSYOEAAMNSLLASSRSEOSFTVVQPFYETTPSLHSDPELOD 300
 Db 241 CSSEETRLAVYVQWMSYOEAAMNSLLASSRSEOSFTVVQPFYETTPSLHSDPELOD 300
 QY 301 STTLAHLNMRMMEPAGEKDEPLSVKGRPMKCPQESPYLFSYRNSNYLTRLOKPODKL 360
 Db 301 STTLAHLNMRMMEPAGEKDEPLSVKGRPMKCPQESPYLFSYRNSNYLTRLOKPODKL 360
 QY 361 EYREGAEIRCPDQDPSPTVPTSVHRLKPADINVIGALGDSLTAAGAGSTPGNVLDVLTQ 420
 Db 361 EYREGAEIRCPDQDPSPTVPTSVHRLKPADINVIGALGDSLTAAGAGSTPGNVLDVLTQ 420
 QY 421 YRGLSMSVGDENIGTTLANTLIREFNPSLKGSVGTGKETSNAFLNQAAGGRBEDL 480
 Db 421 YRGLSMSVGDENIGTTLANTLIREFNPSLKGSVGTGKETSNAFLNQAAGGRBEDL 480
 QY 481 PVOARRLVDMKNDTRIHFOEDWKIITLFTIGNDLCDFCNLDVHYSPOQFTDNIGKALDI 540
 Db 481 PVOARRLVDMKNDTRIHFOEDWKIITLFTIGNDLCDFCNLDVHYSPOQFTDNIGKALDI 540
 QY 541 LHAEBVPAFVNLVTVEIVNLRELYQEKVYCPRMILRSICPCVLFKFDNSTELATLIEF 600
 Db 541 LHAEBVPAFVNLVTVEIVNLRELYQEKVYCPRMILRSICPCVLFKFDNSTELATLIEF 600
 QY 601 NKRFQEKTHOLISGRYDTEDEFTVVQPFENVMKTSBEGLPDNSFPADECHEFSKS 660
 Db 601 NKRFQEKTHOLISGRYDTEDEFTVVQPFENVMKTSBEGLPDNSFPADECHEFSKS 660
 QY 661 HSRBAASALMMNMLEPVQOKTTRHKFNKINITCPNOVQPLRTYKNSMOGHGTWLPICDR 720
 Db 661 HSRBAASALMMNMLEPVQOKTTRHKFNKINITCPNOVQPLRTYKNSMOGHGTWLPICDR 720
 QY 721 APSALHPTSVAHARPADIQVVAALGDSLTAAGNGISGRDLPVTTQYRGLSYSAGGDS 780
 Db 721 APSALHPTSVAHARPADIQVVAALGDSLTAAGNGISGRDLPVTTQYRGLSYSAGGDS 780
 QY 781 LENVTTLPILIREPNRLTGYAVGTGANOTNAFLNQAAGKAEIDMSOVOTLMQMKD 840
 Db 781 LENVTTLPILIREPNRLTGYAVGTGANOTNAFLNQAAGKAEIDMSOVOTLMQMKD 840
 QY 841 DHRVNFHEBPKVITVLIGSDCLCYCTDSNLYSAANVVDLRYNALDYLHREVPVNLV 900
 Db 841 DHRVNFHEBPKVITVLIGSDCLCYCTDSNLYSAANVVDLRYNALDYLHREVPVNLV 900
 QY 901 DFLNPTIMRQVPLGNPDKCPVOQASVLCNCVLTILRENSOELARLEAFSRAYSRSMREL 960
 Db 901 DFLNPTIMRQVPLGNPDKCPVOQASVLCNCVLTILRENSOELARLEAFSRAYSRSMREL 960
 QY 961 SGRVYDQEDSVVLQPFQONIQLPVLADGJPDTSFFAPDCIHPQKHSQIARALMTNML 1020
 Db 961 SGRVYDQEDSVVLQPFQONIQLPVLADGJPDTSFFAPDCIHPQKHSQIARALMTNML 1020
 QY 1021 EPLGSKTETDLEBAEMPITCTPQNEPFLTPRNSNNTYPIKPAIEMNGSFLCTEMKASN 1080
 Db 1021 EPLGSKTETDLEBAEMPITCTPQNEPFLTPRNSNNTYPIKPAIEMNGSFLCTEMKASN 1080
 QY 1081 SVFTSVHQLRPADIKVVAALGDSLTTAVGARPNNSSDLPSTWGLSISIGDGNLEHTHT 1140
 Db 1081 SVFTSVHQLRPADIKVVAALGDSLTTAVGARPNNSSDLPSTWGLSISIGDGNLEHTHT 1140

QY 1141 LPNLLKKNFNYLLGFSTSTWEGTAGLVNAEGARADMPAQANDLVERMKNSPDINLEKD 1200
 Db 1141 LPNLLKKNFNYLLGFSTSTWEGTAGLVNAEGARADMPAQANDLVERMKNSPDINLEKD 1200
 QY 1201 WKLVTLFIVGNDLCHYCENBEAHLATEYVOHIOALDIISEELPRAFVNVVEWELASLY 1260
 Db 1201 WKLVTLFIVGNDLCHYCENBEAHLATEYVOHIOALDIISEELPRAFVNVVEWELASLY 1260
 QY 1261 QGGGKACMLAANNCTCLRHSSQSLKQELKKVNNLQHGISFSFWHOTQREDAV 1320
 Db 1261 QGGGKACMLAANNCTCLRHSSQSLKQELKKVNNLQHGISFSFWHOTQREDAV 1320
 QY 1321 VOPEFQNTLPLNERGDTLTPFSEDCPHFSDRGHAEMAIALMNNMLEPVGRKTSNNFT 1380
 Db 1321 VOPEFQNTLPLNERGDTLTPFSEDCPHFSDRGHAEMAIALMNNMLEPVGRKTSNNFT 1380
 QY 1381 HSRAKLKCPSPESPYLYTLNRSRLPDQAEAEVLYMAVPVAAVGLVVGIICTVVMRC 1440
 Db 1381 HSRAKLKCPSPESPYLYTLNRSRLPDQAEAEVLYMAVPVAAVGLVVGIICTVVMRC 1440
 QY 1441 RRGGRREDPMSLRVAL 1458
 Db 1441 RRGGRREDPMSLRVAL 1458

RESULT 4
 ABU65083
 ID ABU65083 standard; Protein; 1423 AA.
 AC ABU65083;
 XX
 DT 20-MAY-2003 (first entry)
 XX
 DE Human NOV24c protein.
 XX
 KW NOVX; cytosolic; cardiac; antiarteriosclerotic; antiasthmatic; cancer;
 KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
 KW human.
 XX
 OS Homo sapiens.
 XX
 PN MO200272757-A2.
 XX
 PD 19-SEP-2002.
 XX
 PF 08-MAR-2002; 2002MO-US065908.
 XX
 PR 08-MAR-2001; 2001US-274101P.
 PR 08-MAR-2001; 2001US-274194P.
 PR 08-MAR-2001; 2001US-274281P.
 PR 08-MAR-2001; 2001US-274322P.
 PR 09-MAR-2001; 2001US-274849P.
 PR 12-MAR-2001; 2001US-275235P.
 PR 13-MAR-2001; 2001US-275578P.
 PR 13-MAR-2001; 2001US-275579P.
 PR 13-MAR-2001; 2001US-275601P.
 PR 14-MAR-2001; 2001US-276000P.
 PR 16-MAR-2001; 2001US-27676P.
 PR 19-MAR-2001; 2001US-276944P.
 PR 20-MAR-2001; 2001US-277239P.
 PR 20-MAR-2001; 2001US-277321P.
 PR 20-MAR-2001; 2001US-277327P.
 PR 21-MAR-2001; 2001US-277791P.
 PR 22-MAR-2001; 2001US-277833P.
 PR 23-MAR-2001; 2001US-278152P.
 PR 26-MAR-2001; 2001US-278894P.
 PR 27-MAR-2001; 2001US-278999P.
 PR 27-MAR-2001; 2001US-279036P.
 PR 28-MAR-2001; 2001US-279344P.
 PR 30-MAR-2001; 2001US-277338P.
 PR 30-MAR-2001; 2001US-279959P.
 PR 30-MAR-2001; 2001US-280233P.

PR 02-APR-2001; 2001US-280802P.
 PR 02-APR-2001; 2001US-280822P.
 PR 02-APR-2001; 2001US-280900P.
 PR 04-APR-2001; 2001US-281194P.
 PR 13-APR-2001; 2001US-283675P.
 PR 30-APR-2001; 2001US-287424P.
 PR 02-MAY-2001; 2001US-288066P.
 PR 03-MAY-2001; 2001US-288342P.
 PR 03-MAY-2001; 2001US-288528P.
 PR 15-MAY-2001; 2001US-291198P.
 PR 16-MAY-2001; 2001US-291099P.
 PR 30-MAY-2001; 2001US-291240P.
 PR 31-MAY-2001; 2001US-294485P.
 PR 31-MAY-2001; 2001US-294889P.
 PR 18-JUN-2001; 2001US-299027P.
 PR 19-JUN-2001; 2001US-299303P.
 PR 10-JUL-2001; 2001US-299310P.
 PR 10-JUL-2001; 2001US-304354P.
 PR 16-AUG-2001; 2001US-309198P.
 PR 12-SEP-2001; 2001US-318462P.
 PR 12-SEP-2001; 2001US-318770P.
 PR 27-SEP-2001; 2001US-325430P.
 PR 18-OCT-2001; 2001US-325681P.
 PR 31-OCT-2001; 2001US-330380P.
 PR 14-NOV-2001; 2001US-332172P.
 PR 14-NOV-2001; 2001US-332271P.
 PR 14-NOV-2001; 2001US-332272P.
 PR 14-NOV-2001; 2001US-333184P.
 PR 21-NOV-2001; 2001US-332094P.
 PR 03-DEC-2001; 2001US-337426P.
 PR 03-DEC-2001; 2001US-338092P.
 PR 04-DEC-2001; 2001US-337185P.
 PR 03-JAN-2002; 2002US-345705P.
 PR 07-MAR-2002; 2002US-0092900.
 (CURA-) CURAGEN CORP.

PA Padigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CE, Li L;
 PI Zechusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
 PI Paturusan M, Gangoli E, Vernet CAM, Guo X, Tchervnev V;
 PI Fernandez ER, Casman SJ, Malyanar UM, Gerlach V, Liu Y;
 PI Anderson D, Spaderna SK, Catterton E, Burgess C, Lette M, Zhong H;
 PI Alsdbrook JP, Lepley DM, Rieger DK;
 XX WPI; 2002-723332/78.
 DR N-PDB; ABX97050.

PT NOXV polypeptides and polynucleotides, useful for preventing or
 PT treating a disorder associated with aberrant NOXV expression or
 PT activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or
 PT bronchial asthma -

PS Claim 1, Page 170; 1103pp; English.

XX This invention describes novel human NOXV polypeptides which have
 CC cytostatic, cardiant, antiatherosclerotic, antiasthmatic and
 CC hypotensive activity. Pharmaceutical compositions comprising the NOXV
 CC proteins or nucleic acid molecules or NOXV antibodies are useful for
 CC preventing or treating a disorder associated with aberrant NOXV
 CC expression or activity e.g. cancer, hypertension, atherosclerosis,
 CC cardiomyopathy or bronchial asthma. The products of the invention can
 CC be used for gene therapy or in a vaccine. ABU5041-ABU5218 represent
 CC the NOXV polypeptides encoded by ABX97008-ABX97185.

XX Sequence 1423 AA;

Query Match 95.1%; Score 7383.5; DB 23; Length 1423;
 Best Local Similarity 96.4%; Pred. No. 0;
 Matches 1408; Conservative 5; Mismatches 8; Indels 39; Gaps 9;

QY 1 MGLRPGIFILELLLLAGGGTPOIHTSPKSTLEGQLMPEITAKUSPEPCNPNKLGVNMPK 60
 DB 1 MGLRPGIFILELLLLAGGGTPOIHTSPKSTLEGQLMPEITAKUSPEPCNPNKLGVNMPK 40
 QY 61 SVHSLKPSDIKFAVAAGNLEIPDPGTGDEKODWTERPOQVCMGVMTVLSDIIRYSPS 120
 DB 41 -VHSLKPSDIKFAVAAGNLEIPDPGTGDEKOD--ERPQOVCMGVMTVLSDIIRYSPS 97
 QY 121 VPMVCHTGKRVIPHGAEDLWIOAELVNMKENQLOPQPMKLNVEFSNASQCYLC 180
 DB 98 VPMVCHTGKRVIPHGAEDLWIOAELVNMKEN--QLOPQPMKLNVEFSNASQCYLC 156
 QY 181 PSAQONGLAAGVDELMGVLYDQOEVPFAFVNLSEVAEVSROYHGTWLSAPPEPCN 240
 DB 157 PSAQONGLAAGVDELMGVLYDQOEVPFAFVNLSEVAEVSROYHGTWLSAPPEPCN 216
 QY 241 CSEETTRAKVVMQMSYOEAMNSLASRYSEGESFTVVPFPFYEETPSLHSDPRLQD 300
 DB 217 CSEETTRAKVVMQMSYOEAMNSLASRYSEGESFTVVPFPFYEETPS----DPRLOD 272
 QY 301 STTLAMHLNMRMMEPAGKDEPLSYGGRPMKPSQESPTLFESYRNSNYLTRLOKPODKL 360
 DB 273 STTLAMHLNMRMMEPAGKDEPLSYGGRPMKPSQESPTLFESYRNSNYLTRLOKPODKL 332
 QY 361 EVREGAEIRCPDPSDPTVPSVHRLKPADINVIGALGDSLTGNGAGSTPGNVLDVLTQ 420
 DB 333 EVREGAEIRCPDPSDPTVPSVHRLKPADINVIGALGDSLTGNGAGSTPGNVLDVLTQ 392
 QY 421 YRGLSMVSGDENIGVTTLANILREFNPSLKGFSVGTGKETS9PNAFLNOAVAGRAEDL 480
 DB 393 YRGLSMVSGDENIGVTTLANILREFNPSLKGFSVGTGKETS9PNAFLNOAVAGRAEDL 450
 QY 481 PVQARLVLMKNDTRIHPQEDWKITTLFIGNDLDCPNDLVHYS9PONTNDIIGRALDI 540
 DB 451 --QARLVLMKNDTRIHPQEDWKITTLFIGNDLDCPNDLVHYS9PONTNDIIGRALDI 508
 QY 541 LHAEPRAFVNLTVLEIVNLRELIOBKYYCPRMILRSICPVLKFPDDSTELATLIEF 600
 DB 509 LHAEPRAFVNLTVLEIVNLRELIOBKYYCPRMILRSICPVLKFPDDSTELATLIEF 568
 QY 601 NKKEFOETHQILBSGRVDEFTVVVQPFENVMDPKTSEGLPDNSFFAPDCFFHSKKS 660
 DB 569 NKKEFOETHQILBSGRVDEFTVVVQPFENVMDPKTSEGLPDNSFFAPDCFFHSKKS 628
 QY 661 HSRGAASALNMNLEPVGQKTRKFKENKINITCPNOVQ--PLFTYKNSMGCHGTWLP 719
 DB 629 HSRGAASALNMNLEPVGQKTRKFKENKINITCPNOVQ--PLFTYKNSMGCHGTWLP 688
 QY 720 RAPSALHPTSVAHLRPADIOVVAALGSLTAGNGIGSKPDDLDPVTTOYRGLSYSAGDG 779
 DB 689 RAPSALHPTSVAHLRPADIOVVAALGSLTAGNGIGSKPDDLDPVTTOYRGLSYSAGDG 748
 QY 780 SLENVTTLLPILREFNRLTGYAVGTGDANDTNAFLNOAVGAKABDMSQVOTLMQKX 839
 DB 749 SLENVTTLLPILREFNRLTGYAVGTGDANDTNAFLNOAVGAKABDMSQVOTLMQKX 808
 QY 840 DHRVNFHEDEWKITVILIGSDLCYCTDSNLYSAANFVDHLNADLVHREVPRLVNL 899
 DB 809 DHRVNFHEDEWKITVILIGSDLCYCTDSNLYSAANFVDHLNADLVHREVPRLVNL 868
 QY 900 VDLNPTIMQVFLGNPKDPCVQOASVLCVCTLENSQELRLAEFSRAYSSMRELV 959
 DB 869 VDLNPTIMQVFLGNPKDPCVQOASVLCVCTLENSQELRLAEFSRAYSSMRELV 928
 QY 960 GSGRYTQEDSFVYLOPFQONIQLPVLADELPTSEFPAPDCIHPNOKFHSQALRALMTNM 1019
 DB 929 GSGRYTQEDSFVYLOPFQONIQLPVLADELPTSEFPAPDCIHPNOKFHSQALRALMTNM 988
 QY 1020 LEPILSKTETLIDRAEMPTCPTQNEBFLRTPNNSNTYIKALINWGSDFLCTEYKAS 1079
 DB 989 LEPILSKTETLIDRAEMPTCPTQNEBFLRTPNNSNTYIKALINWGSDFLCTEYKAS 1048

QY 1080 NSVPTSVHQLRPADIKVVAALGSLTTAVGARPNNSSDLPTSMRGLSMSIGDGNLETHT 1139
 DB 1049 NSVPTSVHQLRPADIKVVAALGSLTTAVGARPNNSSDLPTSMRGLSMSIGDGNLETHT 1108
 QY 1140 TLNNILKKRPYLLGRTSTSTWEGTAGLVNAAGARADMAQMDLVERKNSP-DINLE 1198
 DB 1109 TLDDILKKRPYLLGRTSTSTWEGTAGLVNAAGARADMAQMDLVERKNSP-DINLE 1168
 QY 1199 KDMKVLTLFIGNVDLCHYCENPAHLATEYVQHIOQLADLISELPRAFNVNVEWELAS 1258
 DB 1169 KDMKVLTLFIGNVDLCHYCENP----VGEYVQHIOQLADLISELPRAFNVNVEWELAS 1224
 QY 1259 LVQGGGKCAMLAANNCTCLRHSSQSLERKQELKKVWNLIQHGISPSYHQYTORDEPA 1318
 DB 1225 LVQGGGKCAMLAANNCTCLRHSSQSLERKQELKKVWNLIQHGISPSYHQYTORDEPA 1284
 QY 1319 VVVQPFQNTLTFLNERGDTLTFFPSDCPHFSDRGAEVAIALMNNMLEPVGRKTTSSN 1378
 DB 1285 VVVQPFQNTLTFLNERGDTLTFFPSDCPHFSDRGAEVAIALMNNMLEPVGRKTTSSN 1343
 QY 1379 FTHSRAKLKCPSPESPYLTTLRNSRLLPDQAEAPETLYWAVPVAAAGVLVGLIGTVW 1438
 DB 1344 FTHSRAKLKCPSPESPYLTTLRNSRLLPDQAEAPETLYWAVPVAAAGVLVGLIGTVW 1403
 QY 1439 RCRGGRREDPPMSLRTVAL 1458
 DB 1404 RCRGGRREDPPMSLRTVAL 1423

RESULT 5
 ABU65081
 ID ABU65081 standard; Protein; 1419 AA.
 XX

AC ABU65081;

DT 20-MAY-2003 (first entry)

DE Human NOV24a protein.

XX NOVA, cytosolic; cardiac; antiarteriosclerotic; antiasthmatic; cancer;
 KW hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
 human.

XX Homo sapiens.

XX MO200272757-A2.

PD 19-SEP-2002.

XX 08-MAR-2002; 2002MO-US06908.

PR 08-MAR-2001; 2001US-274101P.
 PR 08-MAR-2001; 2001US-274194P.
 PR 08-MAR-2001; 2001US-274281P.
 PR 08-MAR-2001; 2001US-274322P.
 PR 09-MAR-2001; 2001US-274849P.
 PR 12-MAR-2001; 2001US-275235P.
 PR 13-MAR-2001; 2001US-275578P.
 PR 13-MAR-2001; 2001US-275579P.
 PR 13-MAR-2001; 2001US-275601P.
 PR 14-MAR-2001; 2001US-276000P.
 PR 16-MAR-2001; 2001US-276776P.
 PR 19-MAR-2001; 2001US-276994P.
 PR 20-MAR-2001; 2001US-277329P.
 PR 20-MAR-2001; 2001US-277327P.
 PR 20-MAR-2001; 2001US-277327P.
 PR 21-MAR-2001; 2001US-277791P.
 PR 22-MAR-2001; 2001US-277833P.
 PR 23-MAR-2001; 2001US-278152P.
 PR 26-MAR-2001; 2001US-278894P.
 PR 27-MAR-2001; 2001US-278999P.
 PR 27-MAR-2001; 2001US-279036P.
 PR 28-MAR-2001; 2001US-279344P.

PR 30-MAR-2001; 2001US-277338P.
 PR 30-MAR-2001; 2001US-279935P.
 PR 30-MAR-2001; 2001US-280233P.
 PR 02-APR-2001; 2001US-280802P.
 PR 02-APR-2001; 2001US-280822P.
 PR 02-APR-2001; 2001US-280900P.
 PR 04-APR-2001; 2001US-281194P.
 PR 13-APR-2001; 2001US-283675P.
 PR 10-APR-2001; 2001US-287424P.
 PR 02-MAY-2001; 2001US-288066P.
 PR 03-MAY-2001; 2001US-288342P.
 PR 03-MAY-2001; 2001US-288528P.
 PR 15-MAY-2001; 2001US-291190P.
 PR 16-MAY-2001; 2001US-291099P.
 PR 16-MAY-2001; 2001US-291240P.
 PR 30-MAY-2001; 2001US-294485P.
 PR 31-MAY-2001; 2001US-294889P.
 PR 31-MAY-2001; 2001US-294899P.
 PR 18-JUN-2001; 2001US-299027P.
 PR 19-JUN-2001; 2001US-299303P.
 PR 10-JUL-2001; 2001US-304354P.
 PR 31-JUL-2001; 2001US-309198P.
 PR 16-AUG-2001; 2001US-312903P.
 PR 10-SEP-2001; 2001US-318462P.
 PR 12-SEP-2001; 2001US-318770P.
 PR 27-SEP-2001; 2001US-325430P.
 PR 27-SEP-2001; 2001US-325681P.
 PR 18-OCT-2001; 2001US-330380P.
 PR 31-OCT-2001; 2001US-335301P.
 PR 14-NOV-2001; 2001US-332172P.
 PR 14-NOV-2001; 2001US-332271P.
 PR 14-NOV-2001; 2001US-332272P.
 PR 14-NOV-2001; 2001US-333184P.
 PR 14-NOV-2001; 2001US-333272P.
 PR 21-NOV-2001; 2001US-332094P.
 PR 03-DEC-2001; 2001US-337426P.
 PR 03-DEC-2001; 2001US-338092P.
 PR 04-DEC-2001; 2001US-337185P.
 PR 03-JAN-2002; 2002US-345705P.
 PR 07-MAR-2002; 2002US-0092900.

(CURA-) CURAGEN CORP.

PA Radigaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CE, Li L;
 XX Zethusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
 PI Paturajan M, Gangoli E, Vernet CM, Guo X, Tcherev V;
 PI Fernandes ER, Casman SJ, Malvankar UM, Gerlach V, Liu Y;
 PI Anderson D, Spaderna SK, Catterton E, Burgess C, Leite M, Zhong H;
 PI Alsobrook JP, Lepley DW, Rieger DK;

DR WPI; 2002-723332/78.
 DR N-PSDB; ABX97048.

XX NOVA polypeptides and polynucleotides, useful for preventing or
 PT treating a disorder associated with aberrant NOVA expression or
 PT activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or
 PT bronchial asthma
 XX
 PS Claim 1; Page 168-169; 1103pp; English.

XX This invention describes novel human NOVA polypeptides which have
 CC cytostatic, cardiac, antiarteriosclerotic, antiasthmatic and
 CC hypotensive activity. Pharmaceutical compositions comprising the NOVA
 CC proteins or nucleic acid molecules or NOVA antibodies are useful for
 CC preventing or treating a disorder associated with aberrant NOVA
 CC expression or activity e.g. cancer, hypertension, atherosclerosis,
 CC cardiomyopathy or bronchial asthma. The products of the invention can
 CC be used for gene therapy or in a vaccine. ABU65041-ABU65218 represent
 CC the NOVA polypeptides encoded by ABX97008-ABX97185.
 XX
 SQ Sequence 1419 AA;

DR WPI; 2002-426287/45.
 DR N-PSDB; AAD37410.
 XX
 PT New human phospholipase-like enzyme polypeptide useful for screening
 PT agents, and in the treatment of cancer, inflammation, diabetes,
 PT obesity, a central nervous system disorder, or a cardiovascular
 PT disorder
 XX
 PS Claim 25, Fig 2; 144pp; English.
 XX
 CC The present invention relates to novel human phospholipase-like enzymes
 CC and polynucleotides encoding such proteins. Sequences of the invention
 CC are useful for producing a medicament for modulating the activity of
 CC phospholipase in a disease such as cancer, inflammation, cardiovascular
 CC disorders, chronic obstructive pulmonary diseases, central nervous system
 CC (CNS) disorders such as brain injuries, cerebrovascular disease, dementia
 CC (Alzheimer's disease), Parkinson's disease, corticobasal degeneration,
 CC motor neuron disease, Pick's disease, Huntington's disease, Creutzfeldt
 CC Jacob dementia, schizophrenia with dementia, Korsakoff's psychosis,
 CC pain associated with CNS (e.g. epilepsy, failed back surgery syndrome,
 CC sciatica), multiple sclerosis, stroke, age associated memory impairment,
 CC allergic disease including asthma, allergic rhinitis (hay fever), atopic
 CC dermatitis, anaphylaxis and inflammation, cardiovascular disease, chronic
 CC obstructive pulmonary disease, acute respiratory distress syndrome, gout,
 CC diabetes, emphysema or obesity. They are also used for treating anorexia,
 CC overweight, cachexia, bulimia, hypertension, type-II diabetes, coronary
 CC artery disease, hyperlipidaemia, gall bladder disease, osteoarthritis,
 CC sleep apnoea and respiratory problems, cancer (e.g. breast, prostate,
 CC colon cancer), thrombolytic disease, reduced fertility, polycystic
 CC ovarian syndrome, complications of pregnancy, menstrual irregularity,
 CC hirsutism, stress incontinence and depression. The present sequence
 CC is human phospholipase like enzyme.
 CC
 XX
 SQ Sequence 1216 AA:
 Query Match 80.5%; Score 6253.5; DB 23; Length 1216;
 Best Local Similarity 97.1%; Pred. No. 0;
 Matches 1181; Conservative 7; Mismatches 9; Indels 19; Gaps 2;
 QY 259 EAMNSILASSRSYSEGSFSTVFPFPFETTPSLHSEDPRLQDSTTLAHLMMNMEPAGE 318
 DB 1 EAMNSILASSRSYSEGSFSTVFPFPFETTPSLHSEDPRLQDSTTLAHLMMNMEPAGE 60
 QY 319 KQEPILSVKGRPMKCSQESPYLFSYRNSNYLRLQKPODKL-----EYRE 364
 DB 61 KQEPILSVKGRPMKCSQESPYLFSYRNSNYLRLQKPODKL-----EYRE 120
 QY 365 GAETRCPPDPSPTVFTSVHRLKPADINVGALGDSLTAGAGSTPGNYLDVLTQYRGL 424
 DB 121 GAETRCPPDPSPTVFTSVHRLKPADINVGALGDSLTAGAGSTPGNYLDVLTQYRGL 180
 QY 425 SMSVGDENIGVTTLANTLIREFNPSLKGFSVGTGKETSNAFLNOAVAGRAEDLPVQA 484
 DB 181 SMSVGDENIGVTTLANTLIREFNPSLKGFSVGTGKETSNAFLNOAVAGRAEDLPVQA 240
 QY 485 RRLVDMKNDTRHFPQDKIITLFTGNDLCPFCDLVHYSQNTDNIQKALDILHAE 544
 DB 241 RRLVDMKNDTRHFPQDKIITLFTGNDLCPFCDLVHYSQNTDNIQKALDILHAE 300
 QY 545 VPRAPFNLTVLEIVNLRELQEKVYCPMILIRSLCPVLKFPDNDSTELATLIEFNKKF 604
 DB 301 VPRAPFNLTVLEIVNLRELQEKVYCPMILIRSLCPVLKFPDNDSTELATLIEFNKKF 360
 QY 605 QEKTHQILIESGRYDREDFTVVQPFPEVNDMPKTSGLPDSNFFAPDCPHFSSKSHSR 664
 DB 361 QEKTHQILIESGRYDREDFTVVQPFPEVNDMPKTSGLPDSNFFAPDCPHFSSKSHSR 420
 QY 665 ASALNNMMLPEVQKTRHFKENKINITCPNOVQPLRTK-----NSMOGHGTWLPQRD 719
 DB 421 ASALNNMMLPEVQKTRHFKENKINITCPNOVQPLRTK-----NSMOGHGTWLPQRD 480
 QY 720 RAPSALHPTSVHALRPADIQVVAALGDSLTPAGNGISGKPPDDLDPVTTYGSLSYSAGDG 779

DB 481 RAPSALHPTSVHALRPADIQVVAALGDSLTPAGNGISGKPPDDLDPVTTYGSLSYSAGDG 540
 QY 780 SLENVTTLNRIIEFNRLTGVAVGTGDANDTNAFLNOAVGAKEDLMSQVOTLQOKM 839
 DB 541 SLENVTTLNRIIEFNRLTGVAVGTGDANDTNAFLNOAVGAKEDLMSQVOTLQOKM 600
 QY 840 DHRVNFHEDMKVITLIGSDLCDYCTDSNLSAANFYDHLNANALDVLHREVPVLVNL 899
 DB 601 DHRVNFHEDMKVITLIGSDLCDYCTDSNLSAANFYDHLNANALDVLHREVPVLVNL 660
 QY 900 VDFLNPITMRQVFLGNPDKCPVOQASVLCNCVLTLENSQELARLEAFSAYSMSRELV 959
 DB 661 VDFLNPITMRQVFLGNPDKCPVOQASVLCNCVLTLENSQELARLEAFSAYSMSRELV 720
 QY 960 GSGRYTOTDPSVYLQPFQNTQLPVLADGLPDTSPAPDCIHPNOKFHSQILRALMTNM 1019
 DB 721 GSGRYTOTDPSVYLQPFQNTQLPVLADGLPDTSPAPDCIHPNOKFHSQILRALMTNM 780
 QY 1020 LEPGSKTETLDRAEMPTICPTQNEPFLTPNSNVTYPIKPAIENWGSDFLCTEWEKAS 1079
 DB 781 LEPGSKTETLDRAEMPTICPTQNEPFLTPNSNVTYPIKPAIENWGSDFLCTEWEKAS 840
 QY 1080 NSVPTSVDLRPADIKVVAALGDSLTTAVGARPNNSSDLPTSWRGLSMSIGDGNLETHT 1139
 DB 841 NSVPTSVDLRPADIKVVAALGDSLTTAVGARPNNSSDLPTSWRGLSMSIGDGNLETHT 900
 QY 1140 TLPNILKKNPYLIGSTSTWECTAGLANVAEGARAPDPAQMDLVERKNSPDITLLEK 1199
 DB 901 TLPNILKKNPYLIGSTSTWECTAGLANVAEGARAPDPAQMDLVERKNSPDITLLEK 960
 QY 1200 DMKLVTLFTIGVNDLCHYCNENPEAHLATEYVOHIOALDILSEELPRAFVNVVEWELASL 1259
 DB 961 DMKLVTLFTIGVNDLCHYCNENPEAHLATEYVOHIOALDILSEELPRAFVNVVEWELASL 1020
 QY 1260 YQGGGKCAMLAAQNNCTCLRHQSLSLEKQELKKNVNNLOHGISPSFYMHTYQREDPAV 1319
 DB 1021 YQGGGKCAMLAAQNNCTCLRHQSLSLEKQELKKNVNNLOHGISPSFYMHTYQREDPAV 1080
 QY 1320 VVQPFQNTLTPLENAGTDLTFPSFEDCFHPSRGAENALALAMNNMLBEVGRKTTSSNNF 1379
 DB 1081 VVQPFQNTLTPLENAGTDLTFPSFEDCFHPSRGAENALALAMNNMLBEVGRKTTSSNNF 1140
 QY 1380 TTSRAKLKCPSPSPYLTLTRNSRLLPDOAEBAPEVLYNAVPVAGVGLVVGIIIGTVVNR 1439
 DB 1141 TTSRAKLKCPSPSPYLTLTRNSRLLPDOAEBAPEVLYNAVPVAGVGLVVGIIIGTVVNR 1200
 QY 1440 CRRGGRREDPMSLRT 1455
 DB 1201 CRRGGRREDPMSLRT 1216
 RESULT 7
 AAM30751
 ID AAM30751 standard; Protein; 1450 AA.
 XX
 AC AAM30751;
 XX
 DT 28-JAN-1998 (first entry)
 XX
 DE Rat phospholipase-B/lipase.
 XX
 KW phospholipase B; lipase; supplement; pancreatic phospholipase; reagent;
 KW screening; rat.
 XX
 OS Rattus rattus.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..30 /label= signal_peptide
 FT Protein 31 /label= mature_protein
 FT Region 43..652 /label= repeat_region_1

Accession	Protein	Location/Qualifiers
14-MAY-2003	(first entry)	
Human	lipid-associated molecule (LIPAM)-9 protein.	
Human	lipid-associated molecule; LIPAM; cardiovascular disorder; str	
arteriovenous fistula; atherosclerosis; hypertension; Raynaud's disea		
aneurysm; congestive heart failure; thrombophlebitis; angina pectoris		
ischemic heart disease; rheumatic heart disease; peptic esophagitis		
gastrointestinal disorder; lipid metabolism disorder; Crohn's disease		
nausea; peptic ulcer; fatty liver; Fabry's disease; Gaucher's disease		
diabetes mellitus; hyperlipidaemia; hypercholesterolaemia; epilepsy;		
autoimmune disorder; inflammatory disorder; neurological disorder; kw		
acquired immunodeficiency syndrome; anaemia; Alzheimer's disease; asst		
dementia; prion disease; Creutzfeldt-Jakob disease; leukaemia; cancer		
adenocarcinoma; lymphoma; melanoma; myeloma; sarcoma; gene therapy;		
protein replacement therapy.		
Homo sapiens.		
Key	Location/Qualifiers	
Peptide	1..19	
/label=	Signal-peptide	
Protein	20..969	
/note=	"Human mature LIPAM-9 protein"	
Peptide	1..21	
/label=	Signal-peptide	
Protein	22..969	
/note=	"Human mature LIPAM-9 protein"	
Peptide	1..22	
/label=	Signal-peptide	
Protein	23..969	
/note=	"Human mature LIPAM-9 protein"	
Peptide	1..25	
/label=	Signal-peptide	
Protein	26..969	
/note=	"Human mature LIPAM-9 protein"	
Domain	393..521	
/note=	"Lipase/acylhydrolase with GDSL-like motif"	
Active-site	394..404	
/note=	"Lipolytic enzyme G-D-S-L family	
serine active-site"		
Domain	740..868	
/note=	"Lipase/acylhydrolase with GDSL-like motif"	
Active-site	741..751	
/note=	"Lipolytic enzyme G-D-S-L family	
serine active-site"		
WO200294988-A2.		
28-NOV-2002.		
17-MAY-2002; 2002WO-US15688.		
18-MAY-2001; 2001US-292242P.		
25-MAY-2001; 2001US-293726P.		
01-JUN-2001; 2001US-295346P.		
06-JUL-2001; 2001US-303404P.		
24-AUG-2001; 2001US-314754P.		
22-JAN-2002; 2002US-351262P.		
29-MAR-2002; 2002US-368799P.		
(INCYTE GENOMICS INC.		
Tang YT, Yue H, Azimzai Y, Baughn MR, Burford N, Reddy R;		
Malina NK, Das D, Nguyen DB, Yao MG, Arvizu CS, Lu Y, Gandhi AR		
Griffin JA, Elliott VS, Ramkumar J, Lai PG, Lu DM, Lee BA;		
Lee SY, Yue H, Yang J, Tribouley CM, Kable AE, Swarnakar A;		
WPI; 2003-120797/11.		
N-PSDB; AAD52634.		
New human lipid-associated molecule (LIPAM) proteins and		

PT polynucleotides, useful for diagnosing, treating or preventing
 PT cardiovascular disorders (e.g. aneurysms), neurological disorders (e.g.
 PT Parkinson's disease) or cancers -
 XX
 XX Claim 64; Page 153-156; 171pp; English.
 CC The present invention relates to novel human lipid-associated molecules
 CC (LIPAM) and polynucleotides encoding such proteins. Sequences of the
 CC invention are useful for treating diseases or conditions associated with
 CC decreased expression of functional LIPAM. The antagonist is useful for
 CC treating a disease or condition associated with the overexpression of
 CC functional LIPAM. They are useful for diagnosing, treating or preventing
 CC cardiovascular disorders (e.g. arteriovenous fistula, atherosclerosis,
 CC hypertension, Raynaud's disease, aneurysms, varicose veins, congestive
 CC heart failure, thrombophlebitis, angina pectoris, ischaemic heart disease
 CC or rheumatic heart disease), gastrointestinal disorders (e.g. peptic
 CC oesophagitis, nausea, peptic ulcer or Crohn's disease), lipid metabolism
 CC disorders (e.g. fatty liver, Fabry's disease, Gaucher's disease, diabetes
 CC mellitus, hyperlipidaemia, hypercholesterolaemia), autoimmune disorders
 CC or inflammatory disorders (e.g. acquired immunodeficiency syndrome,
 CC anaemia, asthma or Crohn's disease), neurological disorders (e.g. stroke,
 CC epilepsy, dementia, Alzheimer's disease), or prion diseases such as kuru
 CC or Creutzfeldt-Jakob disease) or cancers (e.g. adenocarcinoma, leukaemia,
 CC lymphoma, melanoma, myeloma or sarcoma). They are also used in gene
 CC therapy and protein replacement therapy. The present sequence is human
 CC LIPAM-9 protein.
 XX
 XX
 Sequence 969 AA;
 Query Match 63.2%; Score 4911; DB 24; Length 969;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 923; Conservative 0; Mismatches 1; Indels 0; Gaps 0

Qy	1	MGRLPGFLLELLLLLLQGFQOHTSRKSTLEQOLPETLANSPPFCNPKLGVNMSK	60
Dd	1	MGRLPGFLLELLLLLLQGFQOHTSRKSTLEQOLPETLANSPPFCNPKLGVNMSK	60
Qy	61	SVHSLKSDIKFVNAIGNLEIPDPGTDLEKODWTERPQVCVMVTLSDIIRFSPS	120
Dd	61	SVHSLKSDIKFVNAIGNLEIPDPGTDLEKODWTERPQVCVMVTLSDIIRFSPS	120
Qy	121	VPMVCHTGRVIPHDAQEDLMIOAOELVRRMKNLOLDFOPDMKLINVFSSNASQCLC	180
Dd	121	VPMVCHTGRVIPHDAQEDLMIOAOELVRRMKNLOLDFOPDMKLINVFSSNASQCLC	180
Qy	181	PSAONGLAAGGVDELNGVLDYLOOEVPRAVNLVDSEVAEVSROYHGTWLSAPRECN	240
Dd	181	PSAONGLAAGGVDELNGVLDYLOOEVPRAVNLVDSEVAEVSROYHGTWLSAPRECN	240
Qy	241	CSEETTRLAKVMQMSYQOEAUNSLLASSRYSSEQSFYVFOPFYEETTPSLHSEDPRLQD	300
Dd	241	CSEETTRLAKVMQMSYQOEAUNSLLASSRYSSEQSFYVFOPFYEETTPSLHSEDPRLQD	300
Qy	301	STTLAMHLMMRMPEPAGEKDEPLSVKXGRPMKCPSESPLYFSYRNSNYLTRLQKPODXL	360
Dd	301	STTLAMHLMMRMPEPAGEKDEPLSVKXGRPMKCPSESPLYFSYRNSNYLTRLQKPODXL	360
Qy	361	EVRGCAIRCPDCKPSDTPVPSYHRLKPADINVLGALGDSITLAENGASTGCVLDTVLTQ	420
Dd	361	EVRGCAIRCPDCKPSDTPVPSYHRLKPADINVLGALGDSITLAENGASTGCVLDTVLTQ	420
Qy	421	YRGLSMSVGDENGIATYTTLANIILREFNPSLKGFSVGKETSFNPAFLNOAVAGRABDL	480
Dd	421	YRGLSMSVGDENGIATYTTLANIILREFNPSLKGFSVGKETSFNPAFLNOAVAGRABDL	480
Qy	481	PVQARRLVDMKNDTRILHFOEDWKIITLFIGNDLCPFNDVLVHSPONFTDNI GKALDI	540
Dd	481	PVQARRLVDMKNDTRILHFOEDWKIITLFIGNDLCPFNDVLVHSPONFTDNI GKALDI	540
Qy	541	LHAEPVPAFNVLVLEIVNLRELJOEKKYCPMMIIRSLCPCYLKEDDNSTELATLIEF	600
Dd	541	LHAEPVPAFNVLVLEIVNLRELJOEKKYCPMMIIRSLCPCYLKEDDNSTELATLIEF	600

QY 601 NKKFQEKHTQLISGRYDTEDEFTVVVQPFENVDMPTSEGLPDNSFFAPDCFHFSKSKS 660
 DB 601 NKKFQEKHTQLISGRYDTEDEFTVVVQPFENVDMPTSEGLPDNSFFAPDCFHFSKSKS 660
 QY 661 HSRRAASALMNNMLEPVQCKTTRHKFEENKINTCNOVQPLRTYKNSMOGHGWTLPQRDR 720
 DB 661 HSRRAASALMNNMLEPVQCKTTRHKFEENKINTCNOVQPLRTYKNSMOGHGWTLPQRDR 720
 QY 721 APSALHTSVHALRPADIQVVAALGDSLTTAGNGISGRPDDLPTVTTQYRGISYAGSDGS 780
 DB 721 APSALHTSVHALRPADIQVVAALGDSLTTAGNGISGRPDDLPTVTTQYRGISYAGSDGS 780
 QY 781 LENVTTLPNLTRENNMGTGAVGTGDANDTNAFINQAVGAKAEDLMSQVQTLMOQKMD 840
 DB 781 LENVTTLPNLTRENNMGTGAVGTGDANDTNAFINQAVGAKAEDLMSQVQTLMOQKMD 840
 QY 841 DHRVNFHEDMKVITVLLIGSGDLCDYCTDSNLYSANFVDHLRNALDYLAREVPRLVNLV 900
 DB 841 DHRVNFHEDMKVITVLLIGSGDLCDYCTDSNLYSANFVDHLRNALDYLAREVPRLVNLV 900
 QY 901 DFLNPTIMROVFLGNPDKCPVQQA 924
 DB 901 DFLNPTIMROVFLGNPDKCPVQQA 924

RESULT 9

ABP53556
 ID ABP53556 standard; Protein: 472 AA.

AC ABP53556;

DT 16-DEC-2002 (first entry)

DE Human phospholipase protein SEQ ID NO:2.

KW Human; phospholipase; enzyme; chromosome 2.

OS Homo sapiens.

PN NO200262977-A2.

PD 15-AUG-2002.

PF 28-JAN-2002; 2002WO-US02302.

PR 08-FEB-2001; 2001US-0778961.

XX (PEKE) PE CORP NY.

PI Van C, Ketchum KA, Di Francesco V, Beasley EM;

XX WPI; 2002-682698/73.

DR N-PSDB; ABQ82234.

XX New human phospholipase proteins, useful for the development of human
 PT therapeutics and diagnostic compositions, drug screening assays, tissue
 PT typing and pharmacogenomic analysis

PS Claim 1; Fig 2A; 95bp; English.

XX The present sequence represents a human phospholipase protein (I) located
 CC on chromosome 2. (I) can be used for identifying agents that modulate its
 CC function or activity where the agent is useful for treating a disease
 CC or condition mediated by a the human phospholipase protein. (I) peptides
 CC can be used in substantial and specific assays related to functional
 CC information of the peptide sequences, to raise antibodies or to elicit
 CC immune response, as reagents in assays that determine the levels of
 CC protein in biological fluids, and as markers for tissues where the
 CC corresponding protein is expressed. Nucleotide sequences encoding (I)
 CC can be used as probes, primers and chemical intermediates in biological
 CC assays, for constructing recombinant vectors, and expressing antigenic
 CC portions of the protein. (I) and nucleic acid molecules encoding it can
 CC be used in the identification of therapeutic proteins and may serve as

CC models or targets for the development of human therapeutic agents that
 CC modulate phospholipase activity in cells and tissues that express the
 CC phospholipase, such as in kidney, blood, lung, brain glioblastomas,
 CC prostate, colon or leukocytes.

SQ Sequence 472 AA;

Query Match 32.1%; Score 2492; DB 23; Length 472;
 Best local Similarity 93.7%; Pred. No. 2e-216;
 Matches 472; Conservative 0; Mismatches 0; Indels 32; Gaps 1;

QY 955 MRELVSGRYDTEDEFSVVLQPFQNIQLPVLADGLPDTSFAPDCIHQKHSQLARA 1014
 DB 1 MRELVSGRYDTEDEFSVVLQPFQNIQLPVLADGLPDTSFAPDCIHQKHSQLARA 1014
 QY 1015 LMTNMLEPLGSKTETDLRAEMPTICPTONEPPLRTPRNSNTYTPIKPALENMGSDFLCT 1074
 DB 34 -----LEPLGSKTETDLRAEMPTICPTONEPPLRTPRNSNTYTPIKPALENMGSDFLCT 88
 QY 1075 EMKASNSVPTSVHQLRPADIKVVAALGDSLTTAVGARPNNSDLPSTSMRGLSMISGSDGN 1134
 DB 89 EMKASNSVPTSVHQLRPADIKVVAALGDSLTTAVGARPNNSDLPSTSMRGLSMISGSDGN 1134
 QY 1135 LETHHTLPNLTKKFNPYLIGFSTSTWEGTAGLNVALEGARAPMDPAQMDLVERMKNSPD 1194
 DB 149 LETHHTLPNLTKKFNPYLIGFSTSTWEGTAGLNVALEGARAPMDPAQMDLVERMKNSPD 1194
 QY 1195 INLEKMKLVTLFTGVNDLCHYCNENPAHLATRYVQHIOQALDILSEELPRAVNVEVM 1254
 DB 209 INLEKMKLVTLFTGVNDLCHYCNENPAHLATRYVQHIOQALDILSEELPRAVNVEVM 1254
 QY 1255 ELASLVQGGGKCAMLAQNNCTCLRHSQSLERKQELKYNMNIQGISFSYWHQYTOR 1314
 DB 269 ELASLVQGGGKCAMLAQNNCTCLRHSQSLERKQELKYNMNIQGISFSYWHQYTOR 1314
 QY 1315 EDFAVVVQPFQNTLTPLNERRGDTLTFESGDCFHFSDRGHAAEMALMNNMLEPVGRKT 1374
 DB 329 EDFAVVVQPFQNTLTPLNERRGDTLTFESGDCFHFSDRGHAAEMALMNNMLEPVGRKT 1374
 QY 1375 TSNNPHTSRRAKLCPSSESPYLTLTRNSRLPPQAEAPAPVLYWAPVVAAGVGLVGGIIG 1434
 DB 389 TSNNPHTSRRAKLCPSSESPYLTLTRNSRLPPQAEAPAPVLYWAPVVAAGVGLVGGIIG 1434
 QY 1435 TVVWRCRGRGRREDPPMSLRTVAL 1458
 DB 449 TVVWRCRGRGRREDPPMSLRTVAL 1458

RESULT 10

ABU65082
 ID ABU65082 standard; Protein: 310 AA.

AC ABU65082;

DT 20-MAY-2003 (first entry)

DE Human NOV24b protein.

KW NOVX; cytosolic; cardiac; antiarteriosclerotic; antiasthmatic; cancer;
 KM hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
 KM human.

XX Homo sapiens.

PN NO200272757-A2.

PD 19-SEP-2002.

PF 08-MAR-2002; 2002WO-US06908.

XX 08-MAR-2001; 2001US-274101P.

PR 08-MAR-2001; 2001US-274194P.

PR 08-MAR-2001; 2001US-274281P.

PR 08-MAR-2001; 2001US-274322P.
 PR 09-MAR-2001; 2001US-274849P.
 PR 12-MAR-2001; 2001US-275235P.
 PR 13-MAR-2001; 2001US-275578P.
 PR 13-MAR-2001; 2001US-275579P.
 PR 13-MAR-2001; 2001US-275601P.
 PR 14-MAR-2001; 2001US-276000P.
 PR 16-MAR-2001; 2001US-276776P.
 PR 19-MAR-2001; 2001US-276994P.
 PR 20-MAR-2001; 2001US-277239P.
 PR 20-MAR-2001; 2001US-277321P.
 PR 20-MAR-2001; 2001US-277327P.
 PR 21-MAR-2001; 2001US-277791P.
 PR 22-MAR-2001; 2001US-277833P.
 PR 23-MAR-2001; 2001US-278152P.
 PR 26-MAR-2001; 2001US-278894P.
 PR 27-MAR-2001; 2001US-278999P.
 PR 27-MAR-2001; 2001US-279036P.
 PR 28-MAR-2001; 2001US-279344P.
 PR 30-MAR-2001; 2001US-279959P.
 PR 30-MAR-2001; 2001US-280233P.
 PR 02-APR-2001; 2001US-280233P.
 PR 02-APR-2001; 2001US-280802P.
 PR 02-APR-2001; 2001US-280900P.
 PR 04-APR-2001; 2001US-281194P.
 PR 13-APR-2001; 2001US-283675P.
 PR 30-APR-2001; 2001US-287424P.
 PR 02-MAY-2001; 2001US-288066P.
 PR 03-MAY-2001; 2001US-288342P.
 PR 03-MAY-2001; 2001US-288528P.
 PR 15-MAY-2001; 2001US-291190P.
 PR 16-MAY-2001; 2001US-291099P.
 PR 16-MAY-2001; 2001US-291240P.
 PR 30-MAY-2001; 2001US-294485P.
 PR 31-MAY-2001; 2001US-294889P.
 PR 18-JUN-2001; 2001US-299027P.
 PR 19-JUN-2001; 2001US-299303P.
 PR 19-JUN-2001; 2001US-299310P.
 PR 10-JUL-2001; 2001US-304354P.
 PR 31-JUL-2001; 2001US-309198P.
 PR 16-AUG-2001; 2001US-312903P.
 PR 10-SEP-2001; 2001US-318462P.
 PR 12-SEP-2001; 2001US-318770P.
 PR 27-SEP-2001; 2001US-325430P.
 PR 27-SEP-2001; 2001US-325681P.
 PR 18-OCT-2001; 2001US-330380P.
 PR 31-OCT-2001; 2001US-335301P.
 PR 14-NOV-2001; 2001US-332172P.
 PR 14-NOV-2001; 2001US-332271P.
 PR 14-NOV-2001; 2001US-332272P.
 PR 14-NOV-2001; 2001US-333184P.
 PR 14-NOV-2001; 2001US-333272P.
 PR 21-NOV-2001; 2001US-332094P.
 PR 03-DEC-2001; 2001US-337426P.
 PR 03-DEC-2001; 2001US-338092P.
 PR 04-DEC-2001; 2001US-337185P.
 PR 03-JAN-2002; 2002US-345705P.
 PR 07-MAR-2002; 2002US-0092900.
 XX
 RA (CURA-) CURAGEN CORP.
 XX
 PI Padigaru M, Szytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L,
 PI Zernusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
 PI Paturajan M, Gangoli E, Vernet CAM, Guo X, Tcherev V;
 PI Fernandez ER, Casman SJ, Malvankar UM, Gerlach V, Liu Y;
 PI Anderson D, Spaderna SJ, Catterton E, Burgess C, Leite M, Zhong H;
 PI Alsdorck JF, Lepley DM, Rieger DK;
 XX
 DR MPI; 2002-723332/78.
 DR N-PSDB; ABX97049.
 XX

PT NOXV polypeptides and polynucleotides, useful for preventing or
 PT treating a disorder associated with aberrant NOXV expression or
 PT activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or
 PT bronchial asthma
 PS Claim 1; Page 169; 1103pp; English.
 XX
 CC This invention describes novel human NOXV polypeptides which have
 CC cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and
 CC hypotensive activity. Pharmaceutical compositions comprising the NOXV
 CC proteins or nucleic acid molecules or NOXV antibodies are useful for
 CC preventing or treating a disorder associated with aberrant NOXV
 CC expression or activity e.g., cancer, hypertension, atherosclerosis,
 CC cardiomyopathy or bronchial asthma. The products of the invention can
 CC be used for gene therapy or in a vaccine. ABU65041-ABU65218 represent
 CC the NOXV polypeptides encoded by ABX97008-ABX97185.
 XX
 SQ Sequence 310 AA;
 Query Match 19.6%; Score 1525; DB 23; Length 310;
 Best Local Similarity 97.3%; Pred. No. 5.2e-129;
 Matches 285; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
 QY 491 MKNDTRIHFOEDWKITLFIIGNDLDFCNDLVHVSPOHFTDNGKALDILHAEPFAFV 550
 DB 1 MKNDTRIHFOEDWKITLFIIGNDLDFCNDLVHVSPOHFTDNGKALDILHAEPFAFV 60
 QY 551 NLVTVEIYNLRELVOEKKYCPRMILRSICPVLFKPDNDSTELATLIEFNKKFOEKTHQ 610
 DB 61 NLVTVEIYNLRELVOEKKYCPRMILRSICPVLFKPDNDSTELATLIEFNKKFOEKTHQ 120
 QY 611 LIESGRYDREPTVVVOPPEFVNDPKTSEGLPDSFPAPDCFHSKSHSRAASALWN 670
 DB 121 LIESGRYDREPTVVVOPPEFVNDPKTSEGLPDSFPAPDCFHSKSHSRAASALWN 180
 QY 671 NMLEPVGQKTRHKFNKINITCPNOVOPFLRTYKNSMOGHGIMLPCRPAPSAALHPTSV 730
 DB 181 NMLEPVGQKTRHKFNKINITCPNOVOPFLRTYKNSMOGHGIMLPCRPAPSAALHPTSV 240
 QY 731 HALRPADIVVAALGDSLTAAGNGISGSKPDDLPTVTYQRLSYAGGDSLEN 783
 DB 241 HALRPADIVVAALGDSLTAAGNGISGSKPDDLPTVTYQRLSYAGGDSLEN 293
 RESULT 11
 ABB11053
 ID ABB11053 standard; peptide; 267 AA.
 XX
 AC ABB11053;
 XX
 DT 11-JAN-2002 (first entry)
 XX
 DE Human phospholipase B homologue, SEQ ID NO:1423.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; chromolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytostatic; osteopathic; vasotonic; cardiant; virocid; antibacterial;
 KW antifungal; vulnery; antidiabetic.
 XX
 OS Homo sapiens.
 XX
 PN WO200157188-A2.
 XX
 PD 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US03800.
 XX PF 03-FEB-2000; 2000US-0496914.
 XX PR 27-APR-2000; 2000US-0560875.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Tang YT, Liu C, Drmanac RT;
 XX N-PSDB; ABA08297.
 XX DR N-PSDB; ABA08297.
 XX PT Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer -
 PS Claim 20; Page 140; 1963pp; English.
 XX
 XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA0825-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoietic regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g. by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a novel human
 CC polypeptide of the invention.
 CC
 XX
 XX
 SQ Sequence 267 AA;

Query Match 17.5%; Score 1360; DB 22; Length 267;
 Best Local Similarity 73.7%; Pred. No. 3,7e-114;
 Matches 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1;

QY 881 LRNAALDVLRREVPRLVNLVDFLNPITMRQVFGNPDKCPVQQAASVLCNCVLTIRENSQ 940
 DB 1 LRNAALDVLRREVPRLVNLVDFLNPITMRQVFGNPDKCPVQQAASVLCNCVLTIRENSQ 940
 QY 941 LARLEAFSRAYRSMRLVSGRYDQEDFSVVLQPFQNIQLPVLADGJDSFAPDC 1000
 DB 45 ----- 44
 QY 1001 IHNOKFHSOLARALMTNMLEPLGSKTETLDLRAEMPITCPTONEPFLATPRNSNYTPI 1060
 DB ----- 44

DB 45 ----- -MLEPLGSKTETLDLRAEMPITCPTONEPFLATPRNSNYTPI 86
 QY 1061 KPAIENMGSDPFLCTEKASNSVPTSVHOLRPADIKVVAALGDSLTTAVGARPNNSDLP 1120
 DB 87 KPAIENMGSDPFLCTEKASNSVPTSVHOLRPADIKVVAALGDSLTTAVGARPNNSDLP 1160
 QY 1121 SMRGLSMSIGDGNLEHTHTLLPNILKKFNPYLLFGSTSTWEGTAGLVAAEGARADMPA 1180
 DB 147 SMRGLSMSIGDGNLEHTHTLLPNILKKFNPYLLFGSTSTWEGTAGLVAAEGARADMPA 1206
 QY 1181 QAMDLYERMKNSPDINLEKDKMKTLYTFGVNDLCHYCENPEALATRYVQHIOQALDILS 1240
 DB 207 QAMDLYERMKNSPDINLEKDKMKTLYTFGVNDLCHYCENPEALATRYVQHIOQALDILS 1266
 QY 1241 E 1241
 DB 267 E 267

RESULT 12
 AAM25824
 ID AAM25824 standard; Protein; 267 AA.
 XX
 XX AAM25824;
 XX
 PT 16-OCT-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:1339.
 XX
 XX Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
 KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
 KW antibacterial; endocrine; cardiant; central nervous system; viricide;
 KW anti-HIV; fungicide; antimutagen; cardiovascular; antinaemic; anaemia;
 KW antipruritic; haemostatic; vulnery; antidiabetic; osteopathic; eczema;
 KW dermatological; antiallergic; antistimulant; antidiabetic; cytostatic;
 KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
 KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
 KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
 KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
 KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
 KW thrombocytopenia; osteoporosis; severe combined immunodeficiency;
 KW allergic rhinitis; diabetes; multiple sclerosis; depression;
 KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
 KW neurological disorder.
 KW
 XX Homo sapiens.
 OS
 XX
 PN WO200153455-A2.
 XX
 PD 26-JUL-2001.
 XX
 XX 22-DEC-2000; 2000MO-US35017.
 PF
 XX 23-DEC-1999; 99US-0471275.
 PR 21-JAN-2000; 2000US-0488725.
 PR 25-APR-2000; 2000US-0552317.
 XX
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Drmanac RT;
 XX N-PSDB; AAM25824.
 XX N-PSDB; AAM25824.
 XX
 PT Isolated human polynucleotides encoding polypeptides, useful for the
 PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
 PT
 XX
 PS Claim 20; Page 277; 1217pp; English.
 XX
 XX AAM25824 encode the human proteins given in AAM25825 to
 CC AAM25963. The proteins can have activities based on the tissues and
 CC cells they are expressed in, such as: antiinflammatory; antirheumatic;
 CC antitubercular; immunosuppressive; antibacterial; endocrine; cardiant;

CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
 CC cardiovascular; antianemic; antiaggregant; haemostatic; vulnerary;
 CC antitumor; osteopathic; dermatological; antiallergic; antisthmatic;
 CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
 CC encoding them can be used in gene therapy, antisense therapy and vaccine
 CC production. The proteins and polynucleotides are useful for screening for
 CC agonists or antagonists of a protein and for the treatment and diagnosis
 CC of disorders associated with the activity of a protein e.g. inflammation,
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
 CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
 CC neurological disorders.

CC
 XX
 SQ Sequence 267 AA;

Query Match 17.5%; Score 1360; DB 22; Length 267;
 Best Local Similarity 73.7%; Pred. No. 3.7e-114;
 Matches 266; Conservative 0; Mismatches 1; Indels 94; Gaps 1;

QY 881 LRNALDVLAREVPVVLVNLVDFLNPITIMROVFLGNPKCPVOQASVLCNCVTLRENSQE 940
 DB 1 LRNALDVLAREVPVVLVNLVDFLNPITIMROVFLGNPKCPVOQASVLCNCVTLRENSQE 44

QY 941 LARLEAFRAIRSSMRELVSGRYDQEDPSVVLQPFQNIQLPVLADGLPDTSPFAPDC 1000
 DB 45 ----- 44

QY 1001 IHNQKHSOLARALMTNMLEPISKTEITLDAEMPTTCPTONEPPLRTPRNSNTYPTI 1060
 DB 45 -----MLEPISKTEITLDAEMPTTCPTONEPPLRTPRNSNTYPTI 86

QY 1061 KPAIEMGSDFLCTEMKASNSVPTSVHQLRPADIKVVAALGDSLTITVAGARPNNSSDLP 1120
 DB 87 KPAIEMGSDFLCTEMKASNSVPTSVHQLRPADIKVVAALGDSLTITVAGARPNNSSDLP 146

QY 1121 SWSGLWSIGGDNLEHTHTLTPNLLKKFNPLLGFSTSTWEGTAGLVAAEGARADMPA 1180
 DB 147 SWSGLWSIGGDNLEHTHTLTPNLLKKFNPLLGFSTSTWEGTAGLVAAEGARADMPA 206

QY 1181 QANDVLRMKNSPDINLEKMKVLTFLGYNDLCHYENBAHLATEYVGHIOQALDILS 1240
 DB 207 QANDVLRMKNSPDINLEKMKVLTFLGYNDLCHYENBAHLATEYVGHIOQALDILS 266

QY 1241 E 1241
 DB 267 E 267

RESULT 13
 ABB71556
 ID ABB71556 standard; Protein; 424 AA.
 AC ABB71556;
 XX
 XX 26-MAR-2002 (first entry)
 DE Drosophila melanogaster polypeptide SEQ ID NO 41460.
 XX
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.

XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EW;
 XX
 DR WPI; 2001-656960/75.
 DR N-PSDB; ABL15659.
 XX
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 41460; 21pp + Sequence Listing; English.

XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA
 CC sequences (ABU1840-ABU16175) and the encoded proteins
 CC (ABU57737-ABU72072).
 CC
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX
 SQ Sequence 424 AA;

Query Match 7.5%; Score 584; DB 22; Length 424;
 Best Local Similarity 38.5%; Pred. No. 2.3e-43;
 Matches 137; Conservative 57; Mismatches 116; Indels 46; Gaps 12;

QY 369 RCPDSDPTVPTSVRLRPADINVLGALGDSLTAGNAGSTPGNVLDTLYRGLSMVS 428
 DB 88 RSPER-----PTSVRLRPADINVLGALGDSLTAGNAGSTPGNVLDTLYRGLSMVS 139

QY 429 GGDENIGVTTLANILBEPNPSLKGFSVGTGKETSBNAPLANOVACGRADLFPQARRLY 488
 DB 140 GGDENIGVTTLANILBEPNPSLKGFSVGTGKETSBNAPLANOVACGRADLFPQARRLY 199

QY 489 DLMKNTRIHPQEDMKITLFLFGNDLCPQNDLVHY-SPOUNT-----NIGALDILHA 543
 DB 200 RRMQRDPVRVMTSDMKLVTFLFGNN--DFTCDICYPPEKTVDMHBRMLTYRYLRD 256

QY 544 EVPRFVNLVTVLEIVNREL-----YOEKKVVCPRMILRSLCPVLKFDNSTE 593
 DB 257 NVPRMLNIVPA---PNLRFLTNLTGLPPICTIS-----TLRECECLM--GKGKG 302

QY 594 LATLIEFNKKFOEKTHQLIESGRYDREDFTVVVQPPFNVDMPKTSGLPDNSFPAPDC 653
 DB 303 LDYLEGIMKRWIAKOEPIANREBFT-ETFTINVQFQFQDPPEPRRSQGTDRFSEDC 361

QY 654 FHFSSKSHRAASALMNNMLPEVQGT--TRKFEKKNITCPNOVQPLRTYKNS 707
 DB 362 FHSORGHAAANSINNNMLELPGHKSQPAHLFE---TRCPSEMRPPLTIRENS 414

RESULT 14
 ABB96111
 ID ABB96111 standard; Protein; 148 AA.
 AC ABB96111;
 XX
 XX 21-JUN-2002 (first entry)
 DE Human testicular antigen SEQ ID NO: 1495.
 XX
 XX Human; testicular antigen; testes; cancer; metastasis; immune disorder;
 KW reproductive system disorder; urinary system disorder; gene therapy;
 KW cardiovascular disorder; respiratory disorder; neurological disorder;

KM Gastrointestinal disease; infection; cytostatic.
XX Homo sapiens.
XX WO200155317-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01329.
XX PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216860.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225265.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0227009.
PR 01-SEP-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.

PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235634.
PR 27-SEP-2000; 2000US-0235836.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251038.
PR 05-DEC-2000; 2000US-0251038.
PR 05-DEC-2000; 2000US-0251198.
PR 05-DEC-2000; 2000US-0251479.
PR 06-DEC-2000; 2000US-0251856.
PR 06-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251890.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX
XX PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI: 2001-465570/50.
XX N-PSDB; AAL01390.
XX
XX
XX PT Isolated nucleic acid molecule encoding a reproductive system antigen
XX is used in preventing, treating or ameliorating a medical condition -
XX
XX
XX PS Claim 11: SEQ ID NO 4078; 1297pp + Sequence Listing; English.
XX
XX
XX CC The present invention provides the protein and coding sequences of a
XX number of human reproductive system related antigens. These can be used
XX in the prevention and treatment of reproductive system disorders,
XX including cancer. The present sequence is a protein of the invention.
XX
XX
SQ Sequence 148 AA;

Query Match 7.2%; Score 560; DB 22; Length 148;
Best Local Similarity 77.9%; Pred. No. 4.9e-42;
Matches 113; Conservative 11; Mismatches 17; Indels 4; Gaps 3;

QY 1013 RALMTNMLEPLGSKTEFLDLRAEMPTCPONPPFRTPNNSNYTYPIKPAIENMGSDFL 1072
Db 4 QSLMTNMLEPLGSKTEFLDLRAEMPTCPONPPFRTPNNSNYTYPIKPAIENMGSDFL 63
QY 1073 CTEMKASNSVPTSVHQLRPADIKVAALGDSLTTAVGARPNNSSDLPTSWRG-LSWSIGG 1131
Db 64 CTEMKASNSVPTSVHQLRPADIKVAALGDSLTTAVGARPNNSSDLPTSWRG-LSWSIGG 123
QY 1132 DGNL-ETHHTLPNI--LKKNPYLL 1153
Db 124 MGNLGDSSHHTAOSHSEEVQPLPFWLL 148

Search completed: January 6, 2004, 18:56:23
Job time : 86 secs

XX 01-AUG-2002.
 PD 22-JAN-2002; 2002MO-US01715.
 XX 24-JAN-2001; 2001US-264049P.
 XX (LEXI-) LEXICON GENETICS INC.
 PA Yu X, Miranda M, Turner CA;
 PI WPI; 2002-599797/64.
 DR P-PSDB; ABB09555.
 XX Polynucleotides encoding human lipases that are structurally related to
 PT animal lipases, particularly phospholipase B, useful for drug
 PT screening, diagnosis and in gene therapy of biological disorders
 PS Claim 1; Page 36-37; 44pp; English.
 XX The invention relates to a novel human lipase (NHL; ABB09555, ABB09556)
 CC and to nucleic acids encoding it (ABQ77623, ABQ77624). The NHL has
 CC structural similarity with animal lipases, particularly phospholipase B.
 CC Polynucleotides encoding NHL were obtained using human genomic sequences
 CC in conjunction with human thyroid and brain cDNAs. The NHL gene is
 CC located on chromosome 2, and contains a C/T polymorphism at position
 CC 3553 of the open reading frame (ORF), resulting in an Ala/Val
 CC substitution at position 1318 in the protein. NHL nucleotides and
 CC proteins are useful for treating disorders such as inflammatory or
 CC proliferative disease, infectious disease, clotting disorders, and
 CC cancer. They can also be used in screening for compounds useful in
 CC the treatment of mental, biological or medical disorders, as diagnostic
 CC reagents, in clinical trial monitoring and in cosmetic and nutritional
 CC applications. NHL nucleotides can additionally be used in the detection
 CC of disease-associated mutations, in the analysis of gene expression, for
 CC the recombinant expression of NHL, to generate transgenic animals, in
 CC gene therapy, and as part of ribozyme and/or triple helix sequences
 CC useful in the modulation of NHL gene expression. The present sequence
 CC represents cDNA encoding the Ala 1318 variant of NHL.
 XX
 SQ Sequence 4377 BP; 1085 A; 1217 C; 1155 G; 920 T; 0 other;
 Query Match 100.0%; Score 4377; DB 24; Length 4377;
 Best local Similarity 100.0%; Pred. No. 0;
 Matches 4377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 421 TTGGAATTCAGGCTCAAGAACTGTGAGAAACATGAAAGAACTGCAACTTACTTT 480
 DB 421 TTGGAATTCAGGCTCAAGAACTGTGAGAAACATGAAAGAACTGCAACTTACTTT 480
 QY 481 CAATTTGACTGAAAGCTCATCAATGTGTCTTCAATATGCAAGCCAGTGTACTGTGC 540
 DB 481 CAATTTGACTGAAAGCTCATCAATGTGTCTTCAATATGCAAGCCAGTGTACTGTGC 540
 QY 541 CCTCTGCTCAACAGAAATGGGCTTGGCGCGGCGCTGATATAGCTGATGGGGTCTG 600
 DB 541 CCTCTGCTCAACAGAAATGGGCTTGGCGCGGCGCTGATATAGCTGATGGGGTCTG 600
 QY 601 GACTACCTGAGCAGAGAGTCCCAAGACATTTGTAAACCTGTGAGACTCTGAGATT 660
 DB 601 GACTACCTGAGCAGAGAGTCCCAAGACATTTGTAAACCTGTGAGACTCTGAGATT 660
 QY 661 GCAGAGGCTCTGCTGATGATCAAGGCACTTGGCTCAGCCCTGACCAAGCCCTGTAT 720
 DB 661 GCAGAGGCTCTGCTGATGATCAAGGCACTTGGCTCAGCCCTGACCAAGCCCTGTAT 720
 QY 721 TGCTCAGAGAGAACACACCCGCTGGCCAGGTGTATGCAAGTGTATTACAGAAAGCC 780
 DB 721 TGCTCAGAGAGAACACACCCGCTGGCCAGGTGTATGCAAGTGTATTACAGAAAGCC 780
 QY 781 TGGAAACAGCTCCGCGCTCAGCAGAGTACAGTGAAGAGAGTCTTACCGTGTTC 840
 DB 781 TGGAAACAGCTCCGCGCTCAGCAGAGTACAGTGAAGAGAGTCTTACCGTGTTC 840
 QY 841 CAGCTTTCTTATGAGAACACCCCACTCTCACTGAGAGAACCCCACTTCAGAGAT 900
 DB 841 CAGCTTTCTTATGAGAACACCCCACTCTCACTGAGAGAACCCCACTTCAGAGAT 900
 QY 901 TCTAACAGCTGCGCTGCGATCTCTGAAATAGATATGAGACCGAGAGAGAAAGAT 960
 DB 901 TCTAACAGCTGCGCTGCGATCTCTGAAATAGATATGAGACCGAGAGAGAAAGAT 960
 QY 961 GAGCAATGAGTAAACACGGGAGGCAATGAGTCCCTCTCAGAGAGGCGCTAT 1020
 DB 961 GAGCAATGAGTAAACACGGGAGGCAATGAGTCCCTCTCAGAGAGGCGCTAT 1020
 QY 1021 CTGTTCACTTACAGAAACAGCAACTACCTGACAGACTGAGAGAAACCCCAAGACAGCTT 1080
 DB 1021 CTGTTCACTTACAGAAACAGCAACTACCTGACAGACTGAGAGAAACCCCAAGACAGCTT 1080
 QY 1081 GAGGTAAAGAAAGAGCGGAAATCAGATGCTTGAACAAACCCCTCCGATACGTTCC 1140
 DB 1081 GAGGTAAAGAAAGAGCGGAAATCAGATGCTTGAACAAACCCCTCCGATACGTTCC 1140
 QY 1141 ACCTCACTTATAGGCTGAAGCGGCTGACATCAAGTAAATGGAGCCCTGGGTACTCT 1200
 DB 1141 ACCTCACTTATAGGCTGAAGCGGCTGACATCAAGTAAATGGAGCCCTGGGTACTCT 1200
 QY 1201 CTCAGGCAAGCAATGGGCGCGGCTCCACACTGAGAACGCTTGTGACACTGAC 1260
 DB 1201 CTCAGGCAAGCAATGGGCGCGGCTCCACACTGAGAACGCTTGTGACACTGAC 1260
 QY 1261 TACCGAGGCTGTCTGAGCGTGGCGGAGATGAGAACTGGCACTTTACCACTCTG 1320
 DB 1261 TACCGAGGCTGTCTGAGCGTGGCGGAGATGAGAACTGGCACTTTACCACTCTG 1320
 QY 1321 GCGAATCTCTCGGGAATTCAACCCCTCCGTAAGGCGCTCTGTGTGCACTGGGAAA 1380
 DB 1321 GCGAATCTCTCGGGAATTCAACCCCTCCGTAAGGCGCTCTGTGTGCACTGGGAAA 1380
 QY 1381 GAAACAGTCTTATGCTTTTAAACAGGCTGTGGCAGAGAGCCAGCTGAGATCTA 1440
 DB 1381 GAAACAGTCTTATGCTTTTAAACAGGCTGTGGCAGAGAGCCAGCTGAGATCTA 1440
 QY 1441 CTTGTCCAGGCGCAGAGGCTGTGAGACTGATGAAGATGACACAGAGATACCTTTGAG 1500
 DB 1441 CTTGTCCAGGCGCAGAGGCTGTGAGACTGATGAAGATGACACAGAGATACCTTTGAG 1500

QY 1501 GAAGACTGGAAGATAAACCTGTTTATAGCGGCAATGACCTCTGTGATTTCTGCAAT 1560
DB 1501 GAAGACTGGAAGATAAACCTGTTTATAGCGGCAATGACCTCTGTGATTTCTGCAAT 1560
QY 1561 GATCTGCTCACTATTCTCCCGAATCTTACAGACAACTTGGAAAAGCCCTGACATC 1620
DB 1561 GATCTGCTCACTATTCTCCCGAATCTTACAGACAACTTGGAAAAGCCCTGACATC 1620
QY 1621 CTCATGCTGAGGTTCTGGGGCATTTTGGAACTGTGACGGTGTGATGATGTCACAC 1680
DB 1621 CTCATGCTGAGGTTCTGGGGCATTTTGGAACTGTGACGGTGTGATGATGTCACAC 1680
QY 1681 CTGAGGAGCTGTACAGAGAGAAAAAGTCTACTGCGCAAGATGATTCCTCAGGTCTG 1740
DB 1681 CTGAGGAGCTGTACAGAGAGAAAAAGTCTACTGCGCAAGATGATTCCTCAGGTCTG 1740
QY 1741 TGTCCCTGTGTCTTGAAGTTTGAATGATTACTCAACGAACTTGTCTACATGAAATTC 1800
DB 1741 TGTCCCTGTGTCTTGAAGTTTGAATGATTACTCAACGAACTTGTCTACATGAAATTC 1800
QY 1801 AACAAAGATTTCAGAGAGAACCCCACTGATTTGAGAGTGGGGATATGACAAAG 1860
DB 1801 AACAAAGATTTCAGAGAGAACCCCACTGATTTGAGAGTGGGGATATGACAAAG 1860
QY 1861 GAAGATTTTACTGTGTGTGACAGCGTTCTTTGAAAACTGTGACATGCGAAAGACTCG 1920
DB 1861 GAAGATTTTACTGTGTGTGACAGCGTTCTTTGAAAACTGTGACATGCGAAAGACTCG 1920
QY 1921 GAAGATTTGCTGACAACTCTTTCTTGTCTCTGACTGTTTGCATTGACAGCAAGTCT 1980
DB 1921 GAAGATTTGCTGACAACTCTTTCTTGTCTCTGACTGTTTGCATTGACAGCAAGTCT 1980
QY 1981 CACTCCCGAGCAGCGAGTGTCTCTGAAACAATGCTGAGAGCTGTGGCCGAAGAGC 2040
DB 1981 CACTCCCGAGCAGCGAGTGTCTCTGAAACAATGCTGAGAGCTGTGGCCGAAGAGC 2040
QY 2041 ACTCGTCATAGTTTGAAGAAACAGATCAATATCATATGTCGGAACGAGTCCAGCCGTT 2100
DB 2041 ACTCGTCATAGTTTGAAGAAACAGATCAATATCATATGTCGGAACGAGTCCAGCCGTT 2100
QY 2101 CTGAGGACCTTACAGAACAGCATGACAGGTCATGGGACCTGGCTGCATGCAAGGACAG 2160
DB 2101 CTGAGGACCTTACAGAACAGCATGACAGGTCATGGGACCTGGCTGCATGCAAGGACAG 2160
QY 2161 GCGCCCTTCTGCTTGTGACACCTTACTAGTGCATGCGCTGAGACCTGACATCCAAATT 2220
DB 2161 GCGCCCTTCTGCTTGTGACACCTTACTAGTGCATGCGCTGAGACCTGACATCCAAATT 2220
QY 2221 GTGGCTGCTCTGGGGGATTCCTGACCGCTGGCAATGGAAATGGCTCAACAGGACGAC 2280
DB 2221 GTGGCTGCTCTGGGGGATTCCTGACCGCTGGCAATGGAAATGGCTCAACAGGACGAC 2280
QY 2281 CTCGCCGATGTCAACAAGATCGGGGACTGTCAATACAGTGCAGAGGAGGAGCGCTCC 2340
DB 2281 CTCGCCGATGTCAACAAGATCGGGGACTGTCAATACAGTGCAGAGGAGGAGCGCTCC 2340
QY 2341 CTGGAGAAATGTGACCACTTAATCTTCTGCGGAGTTTAAAGAAACCTCAAGGAGC 2400
DB 2341 CTGGAGAAATGTGACCACTTAATCTTCTGCGGAGTTTAAAGAAACCTCAAGGAGC 2400
QY 2401 TAGCGCTGGGACCGGGGTGATGCAATGACAGAAATGCAATCTCAATCAAGCTGTTCC 2460
DB 2401 TAGCGCTGGGACCGGGGTGATGCAATGACAGAAATGCAATCTCAATCAAGCTGTTCC 2460
QY 2461 GGAGCAAGGCTGAGGATCTTATGAGCAAGTCAAACTCTGATGCAAGATGAAAGAT 2520
DB 2461 GGAGCAAGGCTGAGGATCTTATGAGCAAGTCAAACTCTGATGCAAGATGAAAGAT 2520
QY 2521 GATCATAGAGTAAATTTTCAATGAAGCTGAGAGGTCACTCAGTGTCTGATCGAGGAGC 2580
DB 2521 GATCATAGAGTAAATTTTCAATGAAGCTGAGAGGTCACTCAGTGTCTGATCGAGGAGC 2580
QY 2581 GATTATGTGACTACTGACAGATGGAATCTGTATTCTGACAGCAACTTGTGACAT 2640

DB 2581 GATTATGTGACTACTGACAGATGGAATCTGTATTCTGACAGCAACTTGTGACAT 2640
QY 2641 CTCGCCAATGCTTGGAGCTCTTGCATAGAGAGTGGCCAGAGTCTGTGTCAACTCTG 2700
DB 2641 CTCGCCAATGCTTGGAGCTCTTGCATAGAGAGTGGCCAGAGTCTGTGTCAACTCTG 2700
QY 2701 GACTTCTGAACCCCACTATCATGCGGAGGTGTTCTGTGGAAAACCCAGCAAGTCCCA 2760
DB 2701 GACTTCTGAACCCCACTATCATGCGGAGGTGTTCTGTGGAAAACCCAGCAAGTCCCA 2760
QY 2761 GTGACAGAGCCAGCGTTTGTGTAACTGCGTTCTGACCTGTGGAGAACTCCCAAG 2820
DB 2761 GTGACAGAGCCAGCGTTTGTGTAACTGCGTTCTGACCTGTGGAGAACTCCCAAG 2820
QY 2821 CTAGCCAGGCTGAGGCTCTTACAGCGAGCTTACCGGACAGCATGCGCGAGCTGTGGG 2880
DB 2821 CTAGCCAGGCTGAGGCTCTTACAGCGAGCTTACCGGAGAGCATGCGCGAGCTGTGGG 2880
QY 2881 TCAGGCGGCTATGACACGAGAGAGACTTCTGTGTGCTGAGCGCTTCTTCCAGAAC 2940
DB 2881 TCAGGCGGCTATGACACGAGAGAGACTTCTGTGTGCTGAGCGCTTCTTCCAGAAC 2940
QY 2941 ATCCAGCTCCTGTCTGTGGGAGATGGGCTCCAGATACGTCTTCTTGGCCCAAGCTGC 3000
DB 2941 ATCCAGCTCCTGTCTGTGGGAGATGGGCTCCAGATACGTCTTCTTGGCCCAAGCTGC 3000
QY 3001 ATCCAGCCAAATCAGAAATTCATCTCCAGCTGGCCAGAGCCCTTTGACCAATATGCTT 3060
DB 3001 ATCCAGCCAAATCAGAAATTCATCTCCAGCTGGCCAGAGCCCTTTGACCAATATGCTT 3060
QY 3061 GAACCACTTGAAGCAAAACAGAACCTTGAACCTGAGACGAGATGCGCAATCACTGT 3120
DB 3061 GAACCACTTGAAGCAAAACAGAACCTTGAACCTGAGACGAGATGCGCAATCACTGT 3120
QY 3121 CCCACTCAGAAATGAGCCCTTCTGAGAACCCCTCGAAATGATTAACAAGTACCCCATC 3180
DB 3121 CCCACTCAGAAATGAGCCCTTCTGAGAACCCCTCGAAATGATTAACAAGTACCCCATC 3180
QY 3181 AAGCCAGCAATGAGAACTGGGGGAGTGAATCTCTGTGTACAGAGTGAAGGCTTCCAT 3240
DB 3181 AAGCCAGCAATGAGAACTGGGGGAGTGAATCTCTGTGTACAGAGTGAAGGCTTCCAT 3240
QY 3241 AGTGTTCGAACCTCTGTCCAGAGCTCCGACAGAGCAATCAAAATGTGTGGCGGCTG 3300
DB 3241 AGTGTTCGAACCTCTGTCCAGAGCTCCGACAGAGCAATCAAAATGTGTGGCGGCTG 3300
QY 3301 GGTGACTCTCTGACTACAGAGTGGAGCTGACCAAACTCCAGTACCTTACCCACA 3360
DB 3301 GGTGACTCTCTGACTACAGAGTGGAGCTGACCAAACTCCAGTACCTTACCCACA 3360
QY 3361 TCTTGGAGGGAGCTCTTGGAGCATTTGGAGGGAGTGGAACTTGGAGACTCACACACA 3420
DB 3361 TCTTGGAGGGAGCTCTTGGAGCATTTGGAGGGAGTGGAACTTGGAGACTCACACACA 3420
QY 3421 CTGCCCAATTTGTGAAGAGTTTCAACCTTACTCTTGTGCTTCTTACAGCACTGTG 3480
DB 3421 CTGCCCAATTTGTGAAGAGTTTCAACCTTACTCTTGTGCTTCTTACAGCACTGTG 3480
QY 3481 GAGGGGACAGAGAACTTAATGTGACAGGAGGAGGAGGAGCTGAGGAGCATGCGAGCC 3540
DB 3481 GAGGGGACAGAGAACTTAATGTGACAGGAGGAGGAGGAGCTGAGGAGCATGCGAGCC 3540
QY 3541 CAGGCTGTGGACCTGTGAGAGGAAATGAAACAGCCCGGACATCAACTGGAGAAAGAC 3600
DB 3541 CAGGCTGTGGACCTGTGAGAGGAAATGAAACAGCCCGGACATCAACTGGAGAAAGAC 3600
QY 3601 TGAAGCTGTGACACTCTTCAATTTGGGCTCAACGATTTGTGTACTTACTGTGAAATCCG 3660
DB 3601 TGAAGCTGTGACACTCTTCAATTTGGGCTCAACGATTTGTGTACTTACTGTGAAATCCG 3660
QY 3661 GAGGCCCATGAGGACCGGAATATGTTGAGCAATCCACAGGCGCTGACATCTCTCT 3720

Db 3661 GAGGCCACTTGGCCAGCAATATGTTACAGACATCCAGAGCCCTGACATCTCTCT 3720
 Qy 3721 GAGAGCTCCCAAGGGCTTTCATCAAGTGGTGAAGTCAATGAGCTGGCTAGCTCTAC 3780
 Db 3721 GAGAGCTCCCAAGGGCTTTCATCAAGTGGTGAAGTCAATGAGCTGGCTAGCTCTAC 3780
 Qy 3781 CAGGGCCAGAGCGGGAATGTCATGCTGAGCTGAGCAACAACCTGCTGAGCA 3840
 Db 3781 CAGGGCCAGAGCGGGAATGTCATGCTGAGCTGAGCAACAACCTGCTGAGCA 3840
 Qy 3841 CACTCCGAAGCTCTCCGAGAGCAAGAACTGAAGAACTGAAGAACTGGAACCTCAGCAT 3900
 Db 3841 CACTCCGAAGCTCTCCGAGAGCAAGAACTGAAGAACTGAAGAACTGGAACCTCAGCAT 3900
 Qy 3901 GGCATCTCCAGTTTCTCTCTGAGCAATGCAACAGAGGTGAGGCTTTGGGCTTGTG 3960
 Db 3901 GGCATCTCCAGTTTCTCTCTGAGCAATGCAACAGAGGTGAGGCTTTGGGCTTGTG 3960
 Qy 3961 GTCGAGCTTTCTTCCAAAACACACTGACCCCACTGAACGAGAGGGGACACTGACCTC 4020
 Db 3961 GTCGAGCTTTCTTCCAAAACACACTGACCCCACTGAACGAGAGGGGACACTGACCTC 4020
 Qy 4021 ACCTTCTTCTCCGAGAGCTGTTTCACTTCTCAGACCGCGGCAATGCGAGATGCGCATC 4080
 Db 4021 ACCTTCTTCTCCGAGAGCTGTTTCACTTCTCAGACCGCGGCAATGCGAGATGCGCATC 4080
 Qy 4081 GCACCTCGAACAACATGCTGGAACCGAGTGGCGCGCAAGCTCTCTCCAACTTCCACC 4140
 Db 4081 GCACCTCGAACAACATGCTGGAACCGAGTGGCGCGCAAGCTCTCTCCAACTTCCACC 4140
 Qy 4141 CACAGCGGAGCCAACTCAAGTGGCTCTCTGAGAGCCCTTCACTCTACACCTTGGCG 4200
 Db 4141 CACAGCGGAGCCAACTCAAGTGGCTCTCTGAGAGCCCTTCACTCTACACCTTGGCG 4200
 Qy 4201 AACAGCCGATGCTCTCCCAACAGAGCTGAAGAAAGCCCGAGGCTCTTCACTGAGCTGTC 4260
 Db 4201 AACAGCCGATGCTCTCCCAACAGAGCTGAAGAAAGCCCGAGGCTCTTCACTGAGCTGTC 4260
 Qy 4261 CCAGTGGCAGCGGGAAGTGGCTTGTGTTGGGCAATCATCGGGAACAGTGTCTGAGAGTGC 4320
 Db 4261 CCAGTGGCAGCGGGAAGTGGCTTGTGTTGGGCAATCATCGGGAACAGTGTCTGAGAGTGC 4320
 Qy 4321 AGGAGAGTGGCGGAGGAAAGATCTCCATGAGCTGCGACCTGAGCTCTTAG 4377
 Db 4321 AGGAGAGTGGCGGAGGAAAGATCTCCATGAGCTGCGACCTGAGCTCTTAG 4377
 RESULT 2
 ID ABQ77624 standard; cDNA; 4377 BP.
 XX ABQ77624;
 AC 21-OCT-2002 (first entry)
 XX
 DT Human lipase NHL (Val 1318 variant) cDNA.
 XX
 DE Human; lipase; chromosome 2; phospholipase B homologue; thyroid; brain;
 XX inflammatory disease; proliferative disease; infectious disease;
 XX clotting disorder; cancer; drug screening; mental disorder; NHL;
 XX diagnostic reagent; clinical trial monitoring; cosmetic; nutraceutical;
 XX mutation detection; gene expression analysis; transgenic animal;
 XX neoplastic; cytostatic; antiinflammatory; single nucleotide polymorphism;
 KM SNP; gene therapy; gene; 5S.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..4377
 FT /tag= a
 FT /product= "Human lipase NHL"
 FT /replac= (3953, C)
 FT /tag= b
 FT variation

PT /note= "Single nucleotide polymorphism (SNP); leads to
 FT a Val to Ala substitution at position 1318 of
 PT the protein"
 XX
 XX WO200259328-A1.
 XX
 XX 01-AUG-2002.
 XX
 XX 22-JUN-2002; 2002MO-US01715.
 XX
 XX 24-JUN-2001; 2001US-264049P.
 XX
 XX (LEXI-) LEXICON GENETICS INC.
 XX
 XX Yu X, Miranda M, Turner CA,
 XX WPI; 2002-559797/64.
 XX P-PSDB; ABB09556.
 DR
 PT Polynucleotides encoding human lipases that are structurally related to
 PT animal lipases, particularly phospholipase B, useful for drug
 PT screening, diagnosis and in gene therapy of biological disorders
 XX
 XX Disclosure; Page -: 44pp; English.
 XX
 CC The invention relates to a novel human lipase (NHL; ABB09555, ABB09556)
 CC and to nucleic acids encoding it (ABQ77623, ABQ77624). The NHL has
 CC structural similarity with animal lipases, particularly phospholipase B.
 CC Polynucleotides encoding NHL were obtained using human genomic sequences
 CC in conjunction with human thyroid and brain cDNAs. The NHL gene is
 CC located on chromosome 2, and contains a C/T polymorphism at position
 CC 3953 of the open reading frame (ORF), resulting in an Ala/Val
 CC substitution at position 1318 in the protein. NHL nucleotides and
 CC proteins are useful for treating disorders such as inflammatory or
 CC proliferative disease, infectious disease, clotting disorders, and
 CC cancer. They can also be used in screening for compounds useful in
 CC the treatment of mental, biological or medical disorders, as diagnostic
 CC reagents, in clinical trial monitoring and in cosmetic and nutraceutical
 CC applications. NHL nucleotides can additionally be used in the detection
 CC of disease-associated mutations, in the analysis of gene expression, for
 CC the recombinant expression of NHL, to generate transgenic animals, in
 CC gene therapy, and as part of ribozyme and/or triple helix sequences
 CC useful in the modulation of NHL gene expression. The present sequence
 CC represents cDNA encoding the Val 1318 variant of NHL.
 CC Note: The present sequence is not shown in the specification, but
 CC was derived from the information given on page 18 and the Ala 1318
 CC NHL-encoding cDNA (ABQ77623) given in the sequence listing.
 CC
 XX
 XX Sequence 4377 BP; 1085 A; 1216 C; 1155 G; 921 T; 0 other;
 Query Match 100.0%; Score 4375.4; DB 24; Length 4377;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 4376; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 ATGGGGCTGGCGGAGGAGATTTCTCTGAGAGCTCTGCTTCTGGGGCAAGGAGCC 60
 Db 1 ATGGGGCTGGCGGAGGAGATTTCTCTGAGAGCTCTGCTTCTGGGGCAAGGAGCC 60
 Qy 61 CCTCAGATCATATCTCTCTGAGAGAGTATGAGAGGAGGAGCTATGAGGAGGAGCC 120
 Db 61 CCTCAGATCATATCTCTCTGAGAGAGTATGAGAGGAGGAGCTATGAGGAGGAGCC 120
 Qy 121 CTGAGAGATTTCTCATTTCCATGCAACCCAAATTAATTGAGTGAATATGCTTTTAA 180
 Db 121 CTGAGAGATTTCTCATTTCCATGCAACCCAAATTAATTGAGTGAATATGCTTTTAA 180
 Qy 181 TCAGTTCAGCTCTGAGAGCTTCTGATTAATTAATTGAGGAGCCATGAGCAATCTGAA 240
 Db 181 TCAGTTCAGCTCTGAGAGCTTCTGATTAATTAATTGAGGAGCCATGAGCAATCTGAA 240
 Qy 241 ATTCTCTCAGAGCCAGGAGAGCGGCGATCTGAGAGAGCAAGACTGAGCTGAAGGCCAG 300
 Db 241 ATTCTCTCAGAGCCAGGAGAGCGGCGATCTGAGAGAGCAAGACTGAGCTGAAGGCCAG 300

QY 301 CAGGTGATGAGGAGTGAAGACAGTCTTTACAGATCATCATATTTCAAGTCTTCT 360
DB 301 CAGGTGATGAGGAGTGAAGACAGTCTTTACAGATCATCATATTTCAAGTCTTCT 360
QY 361 GTTCCAAATGCTGTGTGCACTGTGAAGAAGACTCATACCCAGATGTGTGAAGAC 420
DB 361 GTTCCAAATGCTGTGTGCACTGTGAAGAAGACTCATACCCAGATGTGTGAAGAC 420
QY 421 TTGTGATTCAGGCTCAAGAACTGTGTGAAGAAGAACTGAAGAAGAACTGTGAAGT 480
DB 421 TTGTGATTCAGGCTCAAGAACTGTGTGAAGAAGAACTGAAGAAGAACTGTGAAGT 480
QY 481 CAATTTGACTGGAAGCTCATCATATGTGTCTTCAATATGAGCCAGTGTACCTGTGC 540
DB 481 CAATTTGACTGGAAGCTCATCATATGTGTCTTCAATATGAGCCAGTGTGTACCTGTGC 540
QY 541 CCTCTGCTCAACAGAAATGGGCTTGGCGGCGGCGGTGATGAGTGTGGGCTGTG 600
DB 541 CCTCTGCTCAACAGAAATGGGCTTGGCGGCGGCGGTGATGAGTGTGGGCTGTG 600
QY 601 GACTTACCTGAGCAGAGAGTCTCCAGAGCATTTGTAACTGTGTGACTTCTGTAGGTT 660
DB 601 GACTTACCTGAGCAGAGAGTCTCCAGAGCATTTGTAACTGTGTGACTTCTGTAGGTT 660
QY 661 GCAGAGGCTCTGTGTGATCAAGGCACTTGGCTCAAGCCCTGACAGAGCCCTGTAT 720
DB 661 GCAGAGGCTCTGTGTGATCAAGGCACTTGGCTCAAGCCCTGACAGAGCCCTGTAT 720
QY 721 TGCTCAGAGGAGACCAACCCGCTGGCCAGAGTGTGTATGAGAGTGTCTTATCAGAAAGC 780
DB 721 TGCTCAGAGGAGACCAACCCGCTGGCCAGAGTGTGTATGAGAGTGTCTTATCAGAAAGC 780
QY 781 TGAACAGGCTCTGTGTGCTCAGAGGATACATGAGCAGAGATCTTCAACCGTGTTC 840
DB 781 TGAACAGGCTCTGTGTGCTCAGAGGATACATGAGCAGAGATCTTCAACCGTGTTC 840
QY 841 CAGCCTTTCTTATGAGACCAACCCCTCTCTCACTCGAGAGGACCCCGACTCCAGAT 900
DB 841 CAGCCTTTCTTATGAGACCAACCCCTCTCTCACTCGAGAGGACCCCGACTCCAGAT 900
QY 901 TCTTACACGCTGTGCTGTGATCTCTGTGAATGAGATGAGAGCAGCAGAGAGAAAGAT 960
DB 901 TCTTACACGCTGTGCTGTGATCTCTGTGAATGAGATGAGAGCAGCAGAGAGAAAGAT 960
QY 961 GAGCCATTGAGTGTAAACACGAGAGGCAATGAAAGTGTCTCTCAGAGAGCCCTAT 1020
DB 961 GAGCCATTGAGTGTAAACACGAGAGGCAATGAAAGTGTCTCTCAGAGAGCCCTAT 1020
QY 1021 CTGTTCAGCTACAGAAACACAGAACTACCTGACAGATGAGAAACCCCAAGACAGCTT 1080
DB 1021 CTGTTCAGCTACAGAAACACAGAACTACCTGACAGATGAGAAACCCCAAGACAGCTT 1080
QY 1081 GAGGTAAAGAAAGAGGAGAAATCAGATGTCTGACAAAGACCCCTCCAGTACGTTCCC 1140
DB 1081 GAGGTAAAGAAAGAGGAGAAATCAGATGTCTGACAAAGACCCCTCCAGTACGTTCCC 1140
QY 1141 ACCTGATTCATAGGCTGAAGCCGCTGACATCAACGTAAATGAGCCCTGAGTGTCT 1200
DB 1141 ACCTGATTCATAGGCTGAAGCCGCTGACATCAACGTAAATGAGCCCTGAGTGTCT 1200
QY 1201 CTGACGAGGAGCAATGGGAGCCGCTGACACCTGTGAGAAAGTCTTGTGACGTCTGAC 1260
DB 1201 CTGACGAGGAGCAATGGGAGCCGCTGACACCTGTGAGAAAGTCTTGTGACGTCTGAC 1260
QY 1261 TACCGAGGCTGTCTGTGAGCCTGTGAGAGATGAGAAACATCGGACCGTTACACCTTG 1320
DB 1261 TACCGAGGCTGTCTGTGAGCCTGTGAGAGATGAGAAACATCGGACCGTTACACCTTG 1320
QY 1321 GCGAATCTCTCGGGAATTCACCTTCTCTGAGAGGCTTCTTGTGCACTGTGGAAA 1380
DB 1321 GCGAATCTCTCGGGAATTCACCTTCTCTGAGAGGCTTCTTGTGCACTGTGGAAA 1380

QY 1381 GAAACAGTCTTATGCTTCTTAAACAGAGCTGTGGCAGAGGCGGAGCTGAGATCTA 1440
DB 1381 GAAACAGTCTTATGCTTCTTAAACAGAGCTGTGGCAGAGGCGGAGCTGAGATCTA 1440
QY 1441 CTTGTCCAGGCGCAGAGGCTGTGTGACCTGTATGAAGATGACACAGAGATACCTTTGAG 1500
DB 1441 CTTGTCCAGGCGCAGAGGCTGTGTGACCTGTATGAAGATGACACAGAGATACCTTTGAG 1500
QY 1501 GAAGCTGGAAGATTAATTAACCCGTTTATAGGGGCAATGACCTGTGTATTTCTGCAAT 1560
DB 1501 GAAGCTGGAAGATTAATTAACCCGTTTATAGGGGCAATGACCTGTGTATTTCTGCAAT 1560
QY 1561 GATCTGATCACTATTCTCCCGAGAACTTCAACAGAACTTGTGAAAAGCCCTGAGCATC 1620
DB 1561 GATCTGATCACTATTCTCCCGAGAACTTCAACAGAACTTGTGAAAAGCCCTGAGCATC 1620
QY 1621 CTCATGCTGAGGCTCTCGGAGATTTGTGAACCTGTGTGACGCTGTGAGATGTCAAC 1680
DB 1621 CTCATGCTGAGGCTCTCGGAGATTTGTGAACCTGTGTGACGCTGTGAGATGTCAAC 1680
QY 1681 CTGAGGAGGCTGTACAGAGAGAAAGAAAGTCTACTGTCCCAAGAGATGATCTCAGGTTCTG 1740
DB 1681 CTGAGGAGGCTGTACAGAGAGAAAGAAAGTCTACTGTCCCAAGAGATGATCTCAGGTTCTG 1740
QY 1741 TGTCCCTGTGCTCTGAAGTTTGTATGATTAACCAAGAACTTGTCTACCTCATGAAATTC 1800
DB 1741 TGTCCCTGTGCTCTGAAGTTTGTATGATTAACCAAGAACTTGTCTACCTCATGAAATTC 1800
QY 1801 AACAGAGGTTTCAGAGAGAAAGACCAACATGATGAGAGTGGGAGATATGACAAAG 1860
DB 1801 AACAGAGGTTTCAGAGAGAAAGACCAACATGATGAGAGTGGGAGATATGACAAAG 1860
QY 1861 GAAGATTTTACTGTGTGTGTGTGACGCGCTTTTGAAGAGTGTGACATGCCAAAGCCTCG 1920
DB 1861 GAAGATTTTACTGTGTGTGTGTGTGACGCGCTTTTGAAGAGTGTGACATGCCAAAGCCTCG 1920
QY 1921 GAAGATTTTACTGTGTGTGTGTGTGACGCGCTTTTGAAGAGTGTGACATGCCAAAGCCTCG 1980
DB 1921 GAAGATTTTACTGTGTGTGTGTGTGACGCGCTTTTGAAGAGTGTGACATGCCAAAGCCTCG 1980
QY 1981 CACTCCGAGCAGCAGTGTCTCTGGAACAAATATGCTGAGAGCCTGTGACAGAAAGC 2040
DB 1981 CACTCCGAGCAGCAGTGTCTCTGGAACAAATATGCTGAGAGCCTGTGACAGAAAGC 2040
QY 2041 ACTGTCTATAGTTTGAAGAAAGATCATATATCATATGTCCGAACAGGTCACGCTTT 2100
DB 2041 ACTGTCTATAGTTTGAAGAAAGATCATATATCATATGTCCGAACAGGTCACGCTTT 2100
QY 2101 CTGAGGACCTTACAGAAACAGCATGAGGATCATGAGACCTGTGCTGACAGAGGACAGA 2160
DB 2101 CTGAGGACCTTACAGAAACAGCATGAGGATCATGAGACCTGTGCTGACAGAGGACAGA 2160
QY 2161 GCCCTTCTGCTTGTGACACCTTACATGATGAGGCTGTGAGACCTGTGAGACATCCAAATT 2220
DB 2161 GCCCTTCTGCTTGTGACACCTTACATGATGAGGCTGTGAGACCTGTGAGACATCCAAATT 2220
QY 2221 GTGGCTGTCTGTGGGAGATTTCTGTGACCGCTGTGCAATGAAATGTGCTCAACAGAGAC 2280
DB 2221 GTGGCTGTCTGTGGGAGATTTCTGTGACCGCTGTGCAATGAAATGTGCTCAACAGAGAC 2280
QY 2281 CTGCGGAGTGTACACACAGATGTGGGAGCTGTCAATGAGTGCAGAGAGGAGGAGCTGC 2340
DB 2281 CTGCGGAGTGTACACACAGATGTGGGAGCTGTCAATGAGTGCAGAGAGGAGGAGCTGC 2340
QY 2341 CTGGAATATGTGACCACTTACTATATCTCTGAGAGTTTAAAGAAACCTCAAGAGC 2400
DB 2341 CTGGAATATGTGACCACTTACTATATCTCTGAGAGTTTAAAGAAACCTCAAGAGC 2400
QY 2401 TACGCGGTGGGACGAGGTATGCAATGACAGAAATGCAATCTCTCAATCAAGCTGTTC 2460
DB 2401 TACGCGGTGGGACGAGGTATGCAATGACAGAAATGCAATCTCTCAATCAAGCTGTTC 2460
QY 2461 GAGAGCAAGGCTGAGAGATCTTATGAGCCAAAGTCCAACTGTATGAGAAAGATGAAGAT 2520

```

Db      2461 GGAGCAAAAGGCTGAGATCTTATGAGCCAAAGTCCAAATCTGATGAGAAAGATGAAGAT 2520
Qy      2521 GATCATAGATAAATTTCCATGAAAGACTGGAAGTCACTACAGTGCCTGATCGAGGACG 2580
Db      2521 GATCATAGATAAATTTCCATGAAAGACTGGAAGTCACTACAGTGCCTGATCGAGGACG 2580
Qy      2581 GATTATAGTACTACTGACAGATTCGAATCTGATCTGAGCAACTTGTGTGACCAT 2640
Db      2581 GATTATAGTACTACTGACAGATTCGAATCTGATCTGAGCAACTTGTGTGACCAT 2640
Qy      2641 CTCGGCAATGCTTGGAGCTCTGCAATAGAGAGGTGCCAAGTCTGTGCAACCTGCTG 2700
Db      2641 CTCGGCAATGCTTGGAGCTCTGCAATAGAGAGGTGCCAAGTCTGTGCAACCTGCTG 2700
Qy      2701 GACTTCTGTAACCCCACTATCATGCGGAGAGTGTCTCTGGGAAACCCAGAACAGTCCCA 2760
Db      2701 GACTTCTGTAACCCCACTATCATGCGGAGAGTGTCTCTGGGAAACCCAGAACAGTCCCA 2760
Qy      2761 GTGAGAGAGGCAAGCCTTTGTGTAACTGGGTTCTGACCTGGGGGAGAACTCCCAAGAG 2820
Db      2761 GTGAGAGAGGCAAGCCTTTGTGTAACTGGGTTCTGACCTGGGGGAGAACTCCCAAGAG 2820
Qy      2821 CTAGCCAGGCTGAGGCTTTCAGCCGAGCCTACCGGAGAGCAATGCGGAGCTGTGGGG 2880
Db      2821 CTAGCCAGGCTGAGGCTTTCAGCCGAGCCTACCGGAGAGCAATGCGGAGCTGTGGGG 2880
Qy      2881 TCAGGCGCTATGACAGCGAGAGAGATCTTCTGTGTGCTGCAAGCCTTCTTCCAGAAC 2940
Db      2881 TCAGGCGCTATGACAGCGAGAGAGATCTTCTGTGTGCTGCAAGCCTTCTTCCAGAAC 2940
Qy      2941 ATCAGCTCCCTGCTGCTGGGGAATGGGCTCCAGATAGTCTTCTTGGCCCAAGCTGC 3000
Db      2941 ATCAGCTCCCTGCTGCTGGGGAATGGGCTCCAGATAGTCTTCTTGGCCCAAGCTGC 3000
Qy      3001 ATCCACCAATGAGAAATTCATCTCCAGCTGGCCAGAGCCTTTGGAACCAATATGCTT 3060
Db      3001 ATCCACCAATGAGAAATTCATCTCCAGCTGGCCAGAGCCTTTGGAACCAATATGCTT 3060
Qy      3061 GAACCACTTGGAAAGCAAAAGAGACCTTGGACTGAGAGAGAGATGCCATATCCTGT 3120
Db      3061 GAACCACTTGGAAAGCAAAAGAGACCTTGGACTGAGAGAGAGATGCCATATCCTGT 3120
Qy      3121 CCACCTGAGATGAGCCTTCTGAGAAACCTCTGGAATGTAATCTACAGTACCCCATC 3180
Db      3121 CCACCTGAGATGAGCCTTCTGAGAAACCTCTGGAATGTAATCTACAGTACCCCATC 3180
Qy      3181 AAGCCAGCATTTGAGAACTGGGCACTGACTTCTCTGTGTAAGAGTGAAGGCTTCCAT 3240
Db      3181 AAGCCAGCATTTGAGAACTGGGCACTGACTTCTCTGTGTAAGAGTGAAGGCTTCCAT 3240
Qy      3241 AGTGTTCGAACCTTCTGTCACAGCTCCGACAGAGAGATCAAAAGTGTGGCCGCTG 3300
Db      3241 AGTGTTCGAACCTTCTGTCACAGCTCCGACAGAGAGATCAAAAGTGTGGCCGCTG 3300
Qy      3301 GGTAACCTCTGACTAGAGAGTGGGAGCTGCAACCAATCTCCATGACTTACCCACA 3360
Db      3301 GGTAACCTCTGACTAGAGAGTGGGAGCTGCAACCAATCTCCATGACTTACCCACA 3360
Qy      3361 TCTTGGAGGGGAACTCTTGGAGCATTTGAGGGAGTGGAACTTGGAGACTCACACACA 3420
Db      3361 TCTTGGAGGGGAACTCTTGGAGCATTTGAGGGAGTGGAACTTGGAGACTCACACACA 3420
Qy      3421 CTGCCCCAATCTTGAAGAAATTCACCTTACCTCTTGGCTTCTTACAGAGCACTGG 3480
Db      3421 CTGCCCCAATCTTGAAGAAATTCACCTTACCTCTTGGCTTCTTACAGAGCACTGG 3480
Qy      3481 GAGGGGACAGAGACTAAATGTGGAGCGGAAAGGGCCAGAGTAAAGGACATGCCAGCC 3540
Db      3481 GAGGGGACAGAGACTAAATGTGGAGCGGAAAGGGCCAGAGTAAAGGACATGCCAGCC 3540
Qy      3541 CAGGCTCTGGAGCTGTGAGAGCAATGAAAAACAGCCCAACATCACTGAGAGAAAGAC 3600

```

```

Db      3541 CAGGCTCTGGAGCTGTGAGAGCAATGAAAAACAGCCCGACATCAACTGAGAAAGAC 3600
Qy      3601 TGGAGCTGGTCACTCTTCAATTTGGGGTCAACGACTTGTGATTAAGTGAATCCG 3660
Db      3601 TGGAGCTGGTCACTCTTCAATTTGGGGTCAACGACTTGTGATTAAGTGAATCCG 3660
Qy      3661 GAGGCCACTTGGCCACGGAATATGTTTCAGCATCAACAGGCCCTGAGCATCTCTCT 3720
Db      3661 GAGGCCACTTGGCCACGGAATATGTTTCAGCATCAACAGGCCCTGAGCATCTCTCT 3720
Qy      3721 GAGAGCTCCCAAGGCTTTCGTAACGTGTGAGAGTCAATGAGTCTGCTAGCTGTAC 3780
Db      3721 GAGAGCTCCCAAGGCTTTCGTAACGTGTGAGAGTCAATGAGTCTGCTAGCTGTAC 3780
Qy      3781 CAGGGCCAAAGGGGGAATGTGCACTGGGAGCTCAGAACCACTGCACTTGGCTCAGA 3840
Db      3781 CAGGGCCAAAGGGGGAATGTGCACTGGGAGCTCAGAACCACTGCACTTGGCTCAGA 3840
Qy      3841 CACTCCCAAGCTCCCTGGAGAAAGCAAGAACTGAAAGTGAACCTGCAAGAT 3900
Db      3841 CACTCCCAAGCTCCCTGGAGAAAGCAAGAACTGAAAGTGAACCTGCAAGAT 3900
Qy      3901 GGCATCTCCAGTTTCTCTTACCTGGCAACCAATACACAGCTGAGAGATTTGGTGTG 3960
Db      3901 GGCATCTCCAGTTTCTCTTACCTGGCAACCAATACACAGCTGAGAGATTTGGTGTG 3960
Qy      3961 GTGCAAGCTTCTTCCAAACACATCACCCTCACTGAAAGAGAGAGGACATGACCTC 4020
Db      3961 GTGCAAGCTTCTTCCAAACACATCACCCTCACTGAAAGAGAGAGGACATGACCTC 4020
Qy      4021 ACCCTCTTCCAGAGACTGTTTCACTTCTCAGACCGGAGGATCCGAGATGGCATC 4080
Db      4021 ACCCTCTTCCAGAGACTGTTTCACTTCTCAGACCGGAGGATCCGAGATGGCATC 4080
Qy      4081 GCACTCTGGAACCAATGCTGGAACCAAGTGGGCGGAGACTTCTCAACCACTTCAAC 4140
Db      4081 GCACTCTGGAACCAATGCTGGAACCAAGTGGGCGGAGACTTCTCAACCACTTCAAC 4140
Qy      4141 CAGAGCGAGCCAACTCAAGTGGCCCTCTCTGAGAGCCCTTACCTTACACCTGGCGG 4200
Db      4141 CAGAGCGAGCCAACTCAAGTGGCCCTCTCTGAGAGCCCTTACCTTACACCTGGCGG 4200
Qy      4201 AACAGCGATTTGCTCCAGACAGGCTGGAAGAACCCCGAGGTGCTTACTGGGCTGC 4260
Db      4201 AACAGCGATTTGCTCCAGACAGGCTGGAAGAACCCCGAGGTGCTTACTGGGCTGC 4260
Qy      4261 CCAAGTGGCAAGCGGAGTGGCTTGTGTGGGATCATGTGGGACAGTGTCTGAGGTGC 4320
Db      4261 CCAAGTGGCAAGCGGAGTGGCTTGTGTGGGATCATGTGGGACAGTGTCTGAGGTGC 4320
Qy      4321 AGGAGAGTGGCCGAGGGAAGATCTTCCAAATGAGCTTGGCACTTGGCCCTCTAG 4377
Db      4321 AGGAGAGTGGCCGAGGGAAGATCTTCCAAATGAGCTTGGCACTTGGCCCTCTAG 4377

```

RESULT 3
 AAD52626
 AAD52626 standard; cDNA; 4607 BP.
 XX
 AAD52626;
 DT 14-MAY-2003 (first entry)
 XX
 DE Human lipid-associated molecule (LIPAM)-1 cDNA.
 XX
 KW Human; lipid-associated molecule; LIPAM; cardiovascular disorder; stroke;
 KW arteriovenous fistula; atherosclerosis; hypertension; Raynaud's disease;
 KW aneurysm; congestive heart failure; thrombophlebitis; angina pectoris;
 KW ischemic heart disease; rheumatic heart disease; peptic esophagitis;
 KW gastrointestinal disorder; lipid metabolism disorder; Crohn's disease;
 KW nausea; peptic ulcer; fatty liver; Fabry's disease; Gaucher's disease;
 KW diabetes mellitus; hyperlipidemia; hypercholesterolemia; epilepsy;
 KW autoimmune disorder; inflammatory disorder; neurological disorder; kuru;

QY 781 TGAACAGCCTCTGGGCTCCAGAGGTACAGTGAAGAGAGTCTTACCGTGTTC 840
DB 789 TGAACAGCCTCTGGGCTCCAGAGGTACAGTGAAGAGAGTCTTACCGTGTTC 848
QY 841 CAGCCTTCTTATGAGACCACTCTCTCACTCGAGAGACCCCGACTTCAGAT 900
DB 849 CAGCCTTCTTATGAGACCACTCTCTCACTCGAGAGACCCCGACTTCAGAT 908
QY 901 TCTAACACGCTGGCTGGCATCTCTGGAATAGATGATGAGCCAGAGAGAAAGAT 960
DB 909 TCTAACACGCTGGCTGGCATCTCTGGAATAGATGATGAGCCAGAGAGAAAGAT 968
QY 961 GAGCATTGAGTAAACAAGGAGGCAATGAAGTCCCTCTCAGAGAGCCCTAT 1020
DB 969 GAGCATTGAGTAAACAAGGAGGCAATGAAGTCCCTCTCAGAGAGCCCTAT 1028
QY 1021 CTGTTCACTTACAGAAACAGCAACTACCTGACAGACTGCAAGAAACCCCAAGCAACTT 1080
DB 1029 CTGTTCACTTACAGAAACAGCAACTACCTGACAGACTGCAAGAAACCCCAAGCAACTT 1088
QY 1081 GAGGTAAAGAGAGAGAGGAGAAATCAGATGTCCTGACAAAGACCCCTCCGATACGTTCC 1140
DB 1089 GAGGTAAAGAGAGAGAGGAGAAATCAGATGTCCTGACAAAGACCCCTCCGATACGTTCC 1148
QY 1141 ACCTCAGTTGATAGGCTGAAGCGGCTGACATCAACGTAATGAGGCTGGGTGACTCT 1200
DB 1149 ACCTCAGTTGATAGGCTGAAGCGGCTGACATCAACGTAATGAGGCTGGGTGACTCT 1208
QY 1201 CTCACGCGCAGGCAATGGGCGCGGTCCACACTGGGAACGCTTTGACGCTTGAACCTG 1260
DB 1209 CTCACGCGCAGGCAATGGGCGCGGTCCACACTGGGAACGCTTTGACGCTTGAACCTG 1268
QY 1261 TACCGAGGCTGTCTGAGAGCTGGCGAGATGAGAACTCGGCACTTACCACTCTG 1320
DB 1269 TACCGAGGCTGTCTGAGAGCTGGCGAGATGAGAACTCGGCACTTACCACTCTG 1328
QY 1321 GCGCAACTCTCGGGGAAATCAACCCCTTCCCTGAAGGCTTCTGTGTCACCTGGAAA 1380
DB 1329 GCGCAACTCTCGGGGAAATCAACCCCTTCCCTGAAGGCTTCTGTGTCACCTGGAAA 1388
QY 1381 GAAACAGTCTTATGCTTCTTAAACAGAGCTGTGCGAGAGCGAGCTGAGATCTA 1440
DB 1389 GAAACAGTCTTATGCTTCTTAAACAGAGCTGTGCGAGAGCGAGCTGAGATCTA 1448
QY 1441 CCGTCCAGGCGCAGAGGCTGTGTGACCTGATGAAGATGACACGAGGATACCTTTCAG 1500
DB 1449 CCGTCCAGGCGCAGAGGCTGTGTGACCTGATGAAGATGACACGAGGATACCTTTCAG 1508
QY 1501 GAAGACTGGAAGATTAATACCTGTTATAGGCGGCAATGACCTGTGATTTCTGCAAT 1560
DB 1509 GAAGACTGGAAGATTAATACCTGTTATAGGCGGCAATGACCTGTGATTTCTGCAAT 1568
QY 1561 GATCTGTGCTCACTATTTCTCCCAAGACTTCAAGACAATTTGGAAGGCTTGGACATC 1620
DB 1569 GATCTGTGCTCACTATTTCTCCCAAGACTTCAAGACAATTTGGAAGGCTTGGACATC 1628
QY 1621 CTCCTAGTGAAGTCTCTCGGCGCAATTTGTGAACCTGATGACGCTGTTGAGATGTCAAC 1680
DB 1629 CTCCTAGTGAAGTCTCTCGGCGCAATTTGTGAACCTGATGACGCTGTTGAGATGTCAAC 1688
QY 1681 CTGAGGAGAGCTGTAACAGAGAAAAAGTCTACGCCCCAAGAGATCTCAAGTCTCTG 1740
DB 1689 CTGAGGAGAGCTGTAACAGAGAAAAAGTCTACGCCCCAAGAGATCTCAAGTCTCTG 1748
QY 1741 TGTCCCTGTGCTGAGAGTTGATGATACCAAGAACTGTGTAACCTTCATCGAATTC 1800
DB 1749 TGTCCCTGTGCTGAGAGTTGATGATACCAAGAACTGTGTAACCTTCATCGAATTC 1808
QY 1801 AACAGAGATTTTCAAGAGAGACCCCAACTGATTTGAGATGAGGCGGATGACAAAG 1860
DB 1809 AACAGAGATTTTCAAGAGAGACCCCAACTGATTTGAGATGAGGCGGATGACAAAG 1868

QY 1861 GAAGATTTTACTGTGTGTGACAGCGCTTCTTTGAAGACGTGACATGCCAAGACTCG 1920
DB 1869 GAAGATTTTACTGTGTGTGACAGCGCTTCTTTGAAGACGTGACATGCCAAGACTCG 1928
QY 1921 GAAGATTTGCTTGACAACTCTTTCTTGCTCTGACTGTTTCCATTGACAGCAAGTCT 1980
DB 1929 GAAGATTTGCTTGACAACTCTTTCTTGCTCTGACTGTTTCCATTGACAGCAAGTCT 1988
QY 1981 CACTCCGAGACCGAGTGTCTCTGGAACAATATGCTGGAAGCTGTGGCCGAAGAGAG 2040
DB 1989 CACTCCGAGACCGAGTGTCTCTGGAACAATATGCTGGAAGCTGTGGCCGAAGAGAG 2048
QY 2041 ACTGCTCATAGTTTGAAGAAACAAGATCAATATCATAGTCCGAACAGAGTCCAGCCGTTT 2100
DB 2049 ACTGCTCATAGTTTGAAGAAACAAGATCAATATCATAGTCCGAACAGAGTCCAGCCGTTT 2108
QY 2101 CTGAGGACCTTACAGAAACAGCATGCAAGGCTGATGGAACCTGCTCCATGAGGACAGA 2160
DB 2109 CTGAGGACCTTACAGAAACAGCATGCAAGGCTGATGGAACCTGCTCCATGAGGACAGA 2168
QY 2161 GCGCCTTCTGCTGACACCTTACCTGAGTGCATGAGCCCTGAGACTGAGCATCCAACTT 2220
DB 2169 GCGCCTTCTGCTGACACCTTACCTGAGTGCATGAGCCCTGAGACTGAGCATCCAACTT 2228
QY 2221 GTGGCTGCTCTGGGGGATTTCTGACCGCTGCAATGGAATTTGGCTCAACAGAGAGAC 2280
DB 2229 GTGGCTGCTCTGGGGGATTTCTGACCGCTGCAATGGAATTTGGCTCAACAGAGAGAC 2288
QY 2281 CTCGCCGATGTCACACAGATATCGGGGACTGTCAATCAGTGCAGAGGGGAGCGCTCC 2340
DB 2289 CTCGCCGATGTCACACAGATATCGGGGACTGTCAATCAGTGCAGAGGGGAGCGCTCC 2348
QY 2341 CTGGAAGATGTGACACCTTAACTTATCTCTGAGAGTTTAAAGAAACCTTCACAGGC 2400
DB 2349 CTGGAAGATGTGACACCTTAACTTATCTCTGAGAGTTTAAAGAAACCTTCACAGGC 2408
QY 2401 TACGCGTGGGCGCAGGCTGATGCAATGACAGAAATGCTCTCAATCAAGCTGTTCC 2460
DB 2409 TACGCGTGGGCGCAGGCTGATGCAATGACAGAAATGCTCTCAATCAAGCTGTTCC 2468
QY 2461 GAGCAAGGCTGAGATCTTATGACCAAGTCCAACTCTGATGCAAGAGATGAAGAT 2520
DB 2469 GAGCAAGGCTGAGATCTTATGACCAAGTCCAACTCTGATGCAAGAGATGAAGAT 2528
QY 2521 GATCATGAGTAATTTCCATGGAAGCTGGAAGCTATCACTGCTGATCGGAGCGAGC 2580
DB 2529 GATCATGAGTAATTTCCATGGAAGCTGGAAGCTATCACTGCTGATCGGAGCGAGC 2588
QY 2581 GATTTATGTACTACTGACAGATGGAATCTGTATTTGACAGCACTTGTGTAACAT 2640
DB 2589 GATTTATGTACTACTGACAGATGGAATCTGTATTTGACAGCACTTGTGTAACAT 2648
QY 2641 CTCGCAATGCTTGAAGCTCTGCAATGAGAGTGGCCAGAGTCTGTGTAACCTGCTG 2700
DB 2649 CTCGCAATGCTTGAAGCTCTGCAATGAGAGTGGCCAGAGTCTGTGTAACCTGCTG 2708
QY 2701 GACTTCTGTAACCCCACTATACGCGGCAAGTGTCTCTGGGAAACCCCAAGAGTCCCA 2760
DB 2709 GACTTCTGTAACCCCACTATACGCGGCAAGTGTCTCTGGGAAACCCCAAGAGTCCCA 2768
QY 2761 GTGAGAGGCGCAGCGTTTGTGTAATGCTGCTGCACTGCGGGGAACTCCCAAGAG 2820
DB 2769 GTGAGAGGCGCAGCGTTTGTGTAATGCTGCTGCACTGCGGGGAACTCCCAAGAG 2828
QY 2821 CTAGCAGGCTGAGAGCTTCAAGCGCTTACCGAGCAGATGCGGAGCTGTGGGG 2880
DB 2829 CTAGCAGGCTGAGAGCTTCAAGCGCTTACCGGAGCAGATGCGGAGCTGTGGGG 2888
QY 2881 TCAAGCGCTTATGACACGAGAGAGACTTCTGTGTGTGCTGACGCTTCTTCAGAAC 2940
DB 2889 TCAAGCGCTTATGACACGAGAGAGACTTCTGTGTGTGCTGACGCTTCTTCAGAAC 2948
QY 2941 ATCCAGCTCCCTGTCTGCGGAGATGGGCTCCAGATACGTCCTTCTTGGCCAGACTGC 3000

Db	2949	ATTCAGCTCCCTGTCTCGCGGATGGGCTCCAGATAGTCTTCTTGCCCGACATGCG	3008
Qy	3001	ATCCACCCAAATCAGAAATTCACCTCCAGCTGGCCAGAGCCCTTTGGACCAATATGCTT	3060
Db	3009	ATCCACCCAAATCAGAAATTCACCTCCAGCTGGCCAGAGCCCTTTGGACCAATATGCTT	3068
Qy	3061	GAACCACTTGGAGCAAAAACAGAGCCCTGGAACCTGAGACGAGAGAGATCCCATCACTGT	3120
Db	3069	GAACCACTTGGAGCAAAAACAGAGCCCTGGAACCTGAGACGAGAGATCCCATCACTGT	3128
Qy	3121	CCCACTGAGATGAGCCCTTCTTGGAACCCCTCGGAATATGAACTACACGTAACCCATC	3180
Db	3129	CCCACTGAGATGAGCCCTTCTTGGAACCCCTCGGAATATGAACTACACGTAACCCATC	3188
Qy	3181	AAGCCAGCCATTGAGAACTGGGGCAGTGACTTCTGTGTACAGATGGAGGCTTCCAA	3240
Db	3189	AAGCCAGCCATTGAGAACTGGGGCAGTGACTTCTGTGTACAGATGGAGGCTTCCAA	3248
Qy	3241	AGTTTCCAACTCTGTCCACCACTGCTCCGACCAAGACATTCAAAGTGTGGCCGCTTG	3300
Db	3249	AGTTTCCAACTCTGTCCACCACTGCTCCGACCAAGACATTCAAAGTGTGGCCGCTTG	3308
Qy	3301	GGTACTCTCTGACCTACAGACAGTGGAGCTGCAACCAACATCTCAGTACCTTCCACA	3360
Db	3309	GGTACTCTCTGACCTACAGACAGTGGAGCTGCAACCAACATCTCAGTACCTTCCACA	3368
Qy	3361	TCTTGGAGGGGACTCTCTTGGAGCATTGGAGGGATGGGAACCTTGGAACTCACACCA	3420
Db	3369	TCTTGGAGGGGACTCTCTTGGAGCATTGGAGGGATGGGAACCTTGGAACTCACACCA	3428
Qy	3421	CTGCCCAACATTTGGAAGAAGTTCAACCTTACTCTTGCGCTTCTTACAGACACTGG	3480
Db	3429	CTGCCCAACATTTGGAAGAAGTTCAACCTTACTCTTGCGCTTCTTACAGACACTGG	3488
Qy	3481	GAGGGGACAGCAGGACTTAAATGTGGACGGGAAGGGGCAAGAGCTAGGGACATGCGACC	3540
Db	3489	GAGGGGACAGCAGGACTTAAATGTGGACGGGAAGGGGCAAGAGCTAGGGACATGCGACC	3548
Qy	3541	CAGGCTTGGAACCTGTATGAGCAGATGAAAAACAGCCCGACATCAACCTTGGAAAAAG	3600
Db	3549	CAGGCTTGGAACCTGTATGAGCAGATGAAAAACAGCCCGACATCAACCTTGGAAAAAG	3608
Qy	3601	TGGAAGCTGGTCACTCTTCTATTGGGGTCAACGACTTGTGTCAATTACTGTGAATCCG	3660
Db	3609	TGGAAGCTGGTCACTCTTCTATTGGGGTCAACGACTTGTGTCAATTACTGTGAATCCG	3668
Qy	3661	GAGGCCCACTTGGACAGGAATATGTTACAGACATTCGAACAGGCCCTGGAAATCTCTCT	3720
Db	3669	GAGGCCCACTTGGACAGGAATATGTTACAGACATTCGAACAGGCCCTGGAAATCTCTCT	3728
Qy	3721	GAGAGCTCCCAAGGGCTTTCGTACACGTGTGGAGGTCACTGAGACTGGCTAGCTGTAC	3780
Db	3729	GAGAGCTCCCAAGGGCTTTCGTACACGTGTGGAGGTCACTGAGACTGGCTAGCTGTAC	3788
Qy	3781	CAGGGCCAAAGCGGGAAAAATGTGCATGTGGCAGCTCGAAACAATGCACTTGGCTCAGA	3840
Db	3789	CAGGGCCAAAGCGGGAAAAATGTGCATGTGGCAGCTCGAAACAATGCACTTGGCTCAGA	3848
Qy	3841	CACCTGGAAAGCTCCCTGTGGAGAACCAAGAAATTGAAGAAAGTGAACCTGCAGAT	3900
Db	3849	CACCTGGAAAGCTCCCTGTGGAGAACCAAGAAATTGAAGAAAGTGAACCTGCAGAT	3908
Qy	3901	GGCATCTCAGATTCTCTTACTAGGACCAATATACACAGCTGTAGAGACTTTCGCGTTGTG	3960
Db	3909	GGCATCTCAGATTCTCTTACTAGGACCAATATACACAGCTGTAGAGACTTTCGCGTTGTG	3968
Qy	3961	GTCGAGCTTTCTTCCAAAAACAACCTACCCCACTGAACGAGAGAGGGGACACTGAACCTC	4020
Db	3969	GTCGAGCTTTCTTCCAAAAACAACCTACCCCACTGAACGAGAGAGGGGACACTGAACCTC	4028
Qy	4021	ACCTTCTTCCGAGGACTGTTTTCACTTTCACACCGCGGGCATGCCGAATGGCCATC	4080

	Db	Accession	Protein Name	Gene	Location/Qualifiers
	Dd	4429	ACCTCTTCTCCAGAGACTGTTTTCACCTTCTCAGAACCGGGGCATGCCAGATTGGCCATC	4088	
Oy	Oy	4081	GCACTCTTGAAACAACATCTGTGGAACACAGTGGGCGGCAAGACTCATCTTCAACCACTTACC	4140	
Dd	Dd	4089	GCACTCTTGAAACAACATCTGTGGAACACAGTGGGCGGCAAGACTCATCTTCAACCACTTACC	4148	
Oy	Oy	4141	CACAGCCGAGCCAAACTCAAAGTGCCCTCTCTGAGAGCCCTTAACCTTACACCTTCGG	4200	
Dd	Dd	4149	CACAGCCGAGCCAAACTCAAAGTGCCCTCTCTGAGAGCCCTTAACCTTACACCTTCGG	4208	
Oy	Oy	4201	AACAGCCGATTTGTCTCCAGACCAAGCTGAAAGAACCCCGAGGTCCTTACTGSGGCTGTG	4260	
Dd	Dd	4209	AACAGCCGATTTGTCTCCAGACCAAGCTGAAAGAACCCCGAGGTCCTTACTGSGGCTGTG	4268	
Oy	Oy	4261	CCAAGTGCACGCGAGTAGTGCGCTTTGTGTGGGCAATCGGAGACAGTGCTGTGAGGTGC	4320	
Dd	Dd	4269	CCAAGTGCACGCGAGTAGTGCGCTTTGTGTGGGCAATCGGAGACAGTGCTGTGAGGTGC	4328	
Oy	Oy	4321	AGGAGAGTGTGCGCGAGAGGAAGATCTCTCAATAGCTGTGGCACTGTGGCCCTTAC	4377	
Dd	Dd	4329	AGGAGAGTGTGCGCGAGAGGAAGATCTCTCAATAGCTGTGGCACTGTGGCCCTTAC	4385	
RESULT 4					
	AAD52634	ID	AAD52634 standard; cDNA; 4424 BP.		
	AAD52634;	XX			
	14-MAY-2003	DT	(first entry)		
	Human lipid-associated molecule (LIPAM) -9 cDNA.	DE			
	Human; lipid-associated molecule; LIPAM; cardiovascular disorder; stroke; arteriovenous fistula; atherosclerosis; hypertension; Raynaud's disease; aneurysm; congestive heart failure; thromboembolism; angina pectoris; ischemic heart disease; rheumatic heart disease; peptic esophagitis; gastrointestinal disorder; lipid metabolism disorder; Crohn's disease; nausea; peptic ulcer; fatty liver; Fabry's disease; Gaucher's disease; diabetes mellitus; hyperlipidaemia; hypercholesterolaemia; epilepsy; autoimmune disorder; inflammatory disorder; neurological disorder; kuru; acquired immunodeficiency syndrome; anaemia; Alzheimer's disease; asthma; dementia; prion disease; Creutzfeldt-Jakob disease; leukaemia; cancer; adenocarcinoma; lymphoma; melanoma; myeloma; sarcoma; gene therapy; protein replacement therapy; gene; ss.	KW			
	Homo sapiens.	XX			
	Key	FH			
	CDS	FT			
	/+tag= a	FT			
	/product= "Human LIPAM-9 protein"	FT			
	9..65	FT			
	/+tag= b	FT			
	66..2915	FT			
	/+tag= c	FT			
	/product= "Human mature LIPAM-9 protein"	FT			
	9..71	FT			
	/+tag= d	FT			
	72..2915	FT			
	/+tag= e	FT			
	/product= "Human mature LIPAM-9 protein"	FT			
	9..74	FT			
	/+tag= f	FT			
	75..2915	FT			
	/+tag= g	FT			
	/product= "Human mature LIPAM-9 protein"	FT			
	9..83	FT			
	/+tag= h	FT			
	84..2915	FT			
	/+tag= i	FT			
	/product= "Human mature LIPAM-9 protein"	FT			
		XX			

XX	PA	(INCY-) INCYTE GENOMICS INC.
XX	PB	Tang Y ⁷ , Yue H, Azimzai Y, Baughn MR, Burford N, Reddy R;
XX	PC	Walla NK, Das D, Nguyen DB, Yao MG, Arvizu CS, Lu Y, Gandhi AR,
XX	PD	Griffin JA, Elliott VS, Ramkumar J, Lal PG, Lu DM, Lee EA;
XX	PE	Lee SY, Yue H, Yang J, Trilbouley CM, Kadle AE, Swarnakar A;
XX	PF	WPI; 2003-120797/11.
XX	PG	P-PSDB; AA634448.
XX	PH	New human lipid-associated molecule (LIPAM) proteins and
XX	PI	polynucleotides, useful for diagnosing, treating or preventing
XX	PJ	cardiovascular disorders (e.g. aneurysms), neurological disorders (e.g.
XX	PK	Parkinson's disease) or cancers -
XX	PL	Claim 74; Page 168-169; 171pp; English.
XX	PM	The present invention relates to novel human lipid-associated molecules
XX	PN	(LIPAM) and polynucleotides encoding such proteins. Sequences of the
XX	PO	invention are useful for treating diseases or conditions associated with
XX	PP	decreased expression of functional LIPAM. The antagonist is useful for
XX	PQ	treating a disease or condition associated with the overexpression of
XX	PR	functional LIPAM. They are useful for diagnosing, treating or preventing
XX	PS	cardiovascular disorders (e.g. arteriovenous fistula, atherosclerosis,
XX	PT	hyper tension, Raynaud's disease, aneurysms, varicose veins, congestive
XX	PV	heart failure, thrombophlebitis, angina pectoris, ischemic heart disease
XX	PW	or rheumatic heart disease), gastrointestinal disorders (e.g. peptic
XX	PX	oesophagitis, nausea, peptic ulcer or Crohn's disease), lipid metabolism
XX	PY	disorders (e.g. fatty liver, Fabry's disease, Gaucher's disease, diabetes
XX	PZ	mellitus, hyperlipidemia, hypercholesterolemia), autoimmune disorders
XX	QA	or inflammatory disorders (e.g. acquired immunodeficiency syndrome,
XX	QB	anaemia, asthma or Crohn's disease), neurological disorders (e.g. stroke,
XX	QC	epilepsy, dementia, Alzheimer's disease, or prion diseases such as kuru
XX	QD	or Creutzfeldt-Jakob disease) or cancers (e.g. adenocarcinoma, leukaemia,
XX	QE	Lymphoma, melanoma, myeloma or sarcoma). They are also used in gene
XX	QF	therapy and protein replacement therapy. The present sequence is human
XX	QG	LIPAM-9 cDNA.
XX	QH	Sequence 4424 BP; 1119 A; 1231 C; 1143 G; 931 T; 0 other;
XX	QI	Query Match 91.0%; Score 3983.4; DB 25; Length 4424;
XX	QJ	Best Local Similarity 95.6%; Pied. No. 0;
XX	QK	Matches 4184; Conservative 0; Mismatches 6; Indels 187; Gaps 1.
XX	QL	1 ATGGGCGTGGGGCCAGGAATTTCTCCTCGGAGCGTCGCGCTTCTGGGGCAAGGACC 60
XX	QM	9 ATGGGCGTGGGGCCAGGAATTTCTCCTCGGAGCGTCGCGCTTCTGGGGCAAGGACC 68
XX	QN	61 CCTCAGATTCACATCCTCTCTGAAGAAGTAGTACATTGGAAGGCAGACTATGCCACAGACC 120
XX	QO	69 CCTCAGATTCACATCCTCTCTGAAGAAGTAGTACATTGGAAGGCAGACTATGCCACAGACC 128
XX	QP	121 CTGAAGAATTCATTCATTCCTCATGCAACCCAAATAAATTGAGAGTAATATGCTTTCTAAA 180
XX	QQ	129 CTGAAGAATTCATTCATTCCTCATGCAACCCAAATAAATTGAGAGTAATATGCTTTCTAAA 188
XX	QR	181 TCAGTTCACCTCTCTGAAGCTTCTGATATTAATTTGTGSGCGCATTTGGCAATCTGGAA 240
XX	QS	189 TCAGTTCACCTCTCTGAAGCTTCTGATATTAATTTGTGSGCGCATTTGGAAATCTGGAA 248

Oy	241	ATTCTCTCAGACCCAGGGAAGGGGCGATCTTGGAAGCAACAACTGACCTGAAAGGCCAG	300
Db	249	ATTCTCTCAGACCCAGGGAAGGGGGAATCTGGAAGCAACAACTGACCTGAAAGGCCAG	308
Oy	301	CAGGTGTGCATGGGAGTGATGACAGTCTCTTACAGACATCATCAGATATTTTCAGTCTTCT	360
Db	309	CAGGTGTGCATGGGAGTGATGACAGTCTCTTACAGACATCATCAGATATTTTCAGTCTTCT	368
Oy	361	GTTCCAAATGCCCTGTGTGCCAOCCTGGAAAGAGTCAATCCCAAGATGGTGTGTAAGAC	420
Db	369	GTTCCAAATGCCCTGTGTGCCACACTGGAAAGAGTCAATCCCAAGATGGTGTGTAAGAC	428
Oy	421	TTTGCGATTACAGGCTCAAGAACTGTGTGAGAAACATGAAAGAAACCTGGAACTTGACTTT	480
Db	429	TTTGCGATTACAGGCTCAAGAACTGTGTGAGAAACATGAAAGAAACCTGGAACTTGACTTT	488
Oy	481	CAATTTAAGTGAAGTCAATCAATGTGTCTTTCAGTAAATCAAGCCAGATTTACTGTGC	540
Db	489	CAATTTAAGTGAAGTCAATCAATGTGTCTTTCAGTAAATCAAGCCAGATTTACTGTGC	548
Oy	541	CCCTTCTCTCAACGAATGGGCTTTGCGGCGGGCGGCTGATGACTGATGGGGGTCTG	600
Db	549	CCCTTCTCTCAACGAATGGGCTTTGCGGCGGGCGGCTGATGACTGATGGGGGTCTG	608
Oy	601	GACTACCTGACAGCAGGAGGTCCCCAGACATTTGTAAACCTGGGTGACCTCTGAGATT	660
Db	609	GACTACCTGACAGCAGGAGGTCCCCAGACATTTGTAAACCTGGGTGACCTCTGAGATT	668
Oy	661	GCAGAGGTCTCTCGTCAGTATCAACGGCACTTGGCTGACCCTGACCAAGACCCTGTAA	720
Db	669	GCAGAGGTCTCTCGTCAGTATCAACGGCACTTGGCTGACCCTGACCAAGACCCTGTAA	728
Oy	721	TGCTCAAGAGAGAACCAACCCGGCTGGCCAAAGTGTGATGCAAGTGTCTTATCAGGAAGCC	780
Db	729	TGCTCAAGAGAGAACCAACCCGGCTGGCCAAAGTGTGATGCAAGTGTCTTATCAGGAAGCC	788
Oy	781	TGGAACAGCCCTCCGGGCTCCAGAGAGGTACAGTGAAGAGAGTCCCTCAACCGATTTTC	840
Db	789	TGGAACAGCCCTCCGGGCTCCAGAGAGGTACAGTGAAGAGAGTCCCTCAACCGATTTTC	848
Oy	841	CAGCCTTCTTCTATGAGACCAACCCCATCTTCACTCGAGAGAACCCCGACCTTCAGAGT	900
Db	849	CAGCCTTCTTCTATGAGACCAACCCCATCTTCACTCGAGAGAACCCCGACCTTCAGAGT	908
Oy	901	TCTACACGCTGGGCTGGCATCTCTGGAATATGATATGATATGAGCCAGAGAGAGAAAGAT	960
Db	909	TCTACACGCTGGGCTGGCATCTCTGGAATATGATATGATATGAGCCAGAGAGAGAAAGAT	968
Oy	961	GAGCCATTGATGTAAAAACACGGGAGGCCAATGAAGTGCCTCTCAGAGAGGCCCTAT	1020
Db	969	GAGCCATTGATGTAAAAACACGGGAGGCCAATGAAGTGCCTCTCAGAGAGGCCCTAT	1028
Oy	1021	CTGTTCAGCTACAGAAAACAGCAACTACCTGACCAAGTGCAGAAAACCCCAAGACAGCTT	1080
Db	1029	CTGTTCAGCTACAGAAAACAGCAACTACCTGACCAAGTGCAGAAAACCCCAAGACAGCTT	1088
Oy	1081	GAGGTAAAGAAAGGAGGGGAAATACAGATGTCCCTGACAAAGACCCCTCGATACGGTTC	1140
Db	1089	GAGGTAAAGAAAGGAGGGGAAATACAGATGTCCCTGACAAAGACCCCTCGATACGGTTC	1148
Oy	1141	ACCTCAGTTCATAGGCTGAAAGCCGGCTGACATCAACTGAAATTTGGAGCCTGGGTGACTCT	1200
Db	1149	ACCTCAGTTCATAGGCTGAAAGCCGGCTGACATCAACTGAAATTTGGAGCCTGGGTGACTCT	1208
Oy	1201	CTCACGGCAGGCAATGGGGCCGGGTCCAACCTTGGGAAACGTTTGGACGTTTACTCTAG	1260
Db	1209	CTCACGGCAGGCAATGGGGCCGGGTCCAACCTTGGGAAACGTTTGGACGTTTACTCTAG	1268
Oy	1261	TACCGAGGCTGTCTGTGAGGCTGTGGGGAGATGAGAAACATCGGCAACGTTTACCAACCTG	1320
Db	1269	TACCGAGGCTGTCTGTGAGGCTGTGGGGAGATGAGAAACATCGGCAACGTTTACCAACCTG	1328

QY 1321 GCGAATCTCTCGGGAATTCACCCCTTCCCTGAAGGGCTTCTGTGGCACTGGGAAA 1380
DB 1329 GCGAATCTCTCGGGAATTCACCCCTTCCCTGAAGGGCTTCTGTGGCACTGGGAAA 1388
QY 1381 GAAACCACTCTTAATGCTTTTAAACCAAGCTGTGGCAGAGGCCGAGCTGAGATCTA 1440
DB 1389 GAAACCACTCTTAATGCTTTTAAACCAAGCTGTGGCAGAGGCCGAGCTGAGATCTA 1448
QY 1441 CCGTGTGAGGCGCAGAGGCTGTGGACCTGATGAAGATGACAAGGATATACCTTTGAG 1500
DB 1449 CCGTGTGAGGCGCAGAGGCTGTGGACCTGATGAAGATGACAAGGATATACCTTTGAG 1508
QY 1501 GAAGACTGGAAGATTAATACCTGTTTATAGGCGGCAATGACCTGTGTGATTTCTGCAAT 1560
DB 1509 GAAGACTGGAAGATTAATACCTGTTTATAGGCGGCAATGACCTGTGTGATTTCTGCAAT 1568
QY 1561 GATCTGTGCTCAATTTCTCCCAAGAACTTCAAGACAACTTGAAGAAAGCCCTGGACATC 1620
DB 1569 GATCTGTGCTCAATTTCTCCCAAGAACTTCAAGACAACTTGAAGAAAGCCCTGGACATC 1628
QY 1621 CTCATGCTGAGGCTCTCGGGGATTTGTGAACCTGTGTGACGGTGTGGATGCTCAAC 1680
DB 1629 CTCATGCTGAGGCTCTCGGGGATTTGTGAACCTGTGTGACGGTGTGGATGCTCAAC 1688
QY 1681 CTGAGGAGCTGTACAGAGAGAAAAAGTCTACTGCCCCAAGATGATCTCAGGCTCTG 1740
DB 1689 CTGAGGAGCTGTACAGAGAGAAAAAGTCTACTGCCCCAAGATGATCTCAGGCTCTG 1748
QY 1741 TGTCCCTGTCTCTGAAGTTTGAATGATTAATCAACAGAACTTGCTGATCTGAAATTC 1800
DB 1749 TGTCCCTGTCTCTGAAGTTTGAATGATTAATCAACAGAACTTGCTGATCTGAAATTC 1808
QY 1801 AACAAGAGTTTGAAG 1860
DB 1809 AACAAGAGTTTGAAG 1868
QY 1861 GAAGATTTTACTGTGTGTGACAGCGCTTCTTGAAGAAAGTGTGACATGCAAGAGCTG 1920
DB 1869 GAAGATTTTACTGTGTGTGACAGCGCTTCTTGAAGAAAGTGTGACATGCAAGAGCTG 1928
QY 1921 GAAGATTTGCTGACAACTCTTTCTTGTGCTGTGACCTGTTTCACTTCAAGAGAGCT 1980
DB 1929 GAAGATTTGCTGACAACTCTTTCTTGTGCTGTGACCTGTTTCACTTCAAGAGAGCT 1988
QY 1981 CACTCCCGAGACCAAGCTGTCTGTGAACAATATGCTGAGAGCTGTGGCCAGAAAGC 2040
DB 1989 CACTCCCGAGACCAAGCTGTCTGTGAACAATATGCTGAGAGCTGTGGCCAGAAAGC 2048
QY 2041 ACTGCTCATTAAGTTTGAAGAGAGATCAATATCAATGCTGAGAGAGAGAGAGAGAG 2100
DB 2049 ACTGCTCATTAAGTTTGAAGAGAGATCAATATCAATGCTGAGAGAGAGAGAGAGAG 2108
QY 2101 CTGAGAGACTTACAAG 2160
DB 2109 CTGAGAGACTTACAAG 2168
QY 2161 GCGCCCTTCTGCTTGGACCTTACTGATGATGCTGTGAGAGCTGTGAGACATCAAGTT 2220
DB 2169 GCGCCCTTCTGCTTGGACCTTACTGATGATGCTGTGAGAGCTGTGAGACATCAAGTT 2228
QY 2221 GTGGGCTGTGGGGGATTTCTGACCGGCTGGCAATGGAATTTGGCTCAAGAGAGAG 2280
DB 2229 GTGGGCTGTGGGGGATTTCTGACCGGCTGGCAATGGAATTTGGCTCAAGAGAGAG 2288
QY 2281 CTCGCCGATGTACCAACAGATATCGGGGACTGTCAATACATGACAGAGAGAGAGAG 2340
DB 2289 CTCGCCGATGTACCAACAGATATCGGGGACTGTCAATACATGACAGAGAGAGAGAG 2348
QY 2341 CTGAGAGATGTACCAACCTTAAATATCTTGGGAGTTTAAACAGAAACCTTCAAGGC 2400
DB 2349 CTGAGAGATGTACCAACCTTAAATATCTTGGGAGTTTAAACAGAAACCTTCAAGGC 2408
QY 2401 TAGCGGTGGGACAGGGTGTATGCAATGACAGAGATCAATTCCTCAATCAAGGTGTTCC 2460

DB 2409 TAGCGGTGGGACAGGGTGTATGCAATGACAGAGATCAATTCCTCAATCAAGGTGTTCC 2468
QY 2461 GAGAGAAAGCTGAGAGATCTTATAGAGCAAGTCCAACTCTGATGAGAGAGAGAGAT 2520
DB 2469 GAGAGAAAGCTGAGAGATCTTATAGAGCAAGTCCAACTCTGATGAGAGAGAGAGAT 2528
QY 2521 GATCATAGAGTAAATTTTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580
DB 2529 GATCATAGAGTAAATTTTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2588
QY 2581 GATTATGTACTACTGACAGAGATGGAATCTGATATTTGACAGCAACTTTGTTGACAT 2640
DB 2589 GATTATGTACTACTGACAGAGATGGAATCTGATATTTGACAGCAACTTTGTTGACAT 2648
QY 2641 CTCGCCGATGCTTGGAGCTCTGCTGATAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2700
DB 2649 CTCGCCGATGCTTGGAGCTCTGCTGATAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2708
QY 2701 GACTTCTGAAACCCCACTATCATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2760
DB 2709 GACTTCTGAAACCCCACTATCATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2768
QY 2761 GTGACAG 2820
DB 2769 GTGACAG 2828
QY 2821 CTAG 2880
DB 2829 CTAG 2888
QY 2881 TCAG 2940
DB 2889 TCAG 2948
QY 2941 ATCCAGCTCCCTGTCTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000
DB 2949 ATCCAGCTCCCTGTCTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3008
QY 3001 ATCCAGCTCCCTGTCTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060
DB 3009 ATCCAGCTCCCTGTCTCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3068
QY 3061 GAACCACTTGAAG 3120
DB 3069 GAACCACTTGAAG 3128
QY 3121 CCCACTCAG 3180
DB 3129 CCCACTCAG 3188
QY 3181 AAGCAG 3240
DB 3189 AAGCAG 3248
QY 3241 AGTGTTCGAACCTTGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3300
DB 3249 AGTGTTCGAACCTTGTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3308
QY 3301 GGTGACTCTCTGACTCAG 3360
DB 3309 GGTGACTCTCTGACTCAG 3368
QY 3361 TCTTGAAG 3420
DB 3369 TCTTGAAG 3428
QY 3421 CTGCGCAATCTTGAAG 3480
DB 3429 CTGCGCAATCTTGAAG 3488
QY 3481 GAGGGAG 3540

Db 3302 GAGGGGACAGCAGACTTAAATGTGGCAGCGAAGGGCCAGAGCTAGAGCATGCCAGCC 3361
 Qy 3541 CAGGCGCTGGGACCTGTAGAGCGAATGAAAAACAGCCCGACATCACTGGAGAAAGAC 3600
 Db 3362 CAGGCGCTGGGACCTGTAGAGCGAATGAAAAACAGCCCGACATCACTGGAGAAAGAC 3421
 Qy 3601 TGGAAAGCTGGTCACTCTTCATTTGGGGTCAACGACTTGTCTATTACTGTAGAAATCCG 3660
 Db 3422 TGGAAAGCTGGTCACTCTTCATTTGGGGTCAACGACTTGTCTATTACTGTAGAAATCCG 3481
 Qy 3661 GAGGCGCTGGGACCTGTAGAGCGAATGAAAAACAGCCCGACATCACTGGAGAAAGAC 3720
 Db 3482 GAGGCGCTGGGACCTGTAGAGCGAATGAAAAACAGCCCGACATCACTGGAGAAAGAC 3541
 Qy 3721 GAGGAGCTCCCAAGGGCTTTCGTCAACGTGGTGGAGGTCATGGAGCTGGAGCTGTAC 3780
 Db 3542 GAGGAGCTCCCAAGGGCTTTCGTCAACGTGGTGGAGGTCATGGAGCTGGAGCTGTAC 3601
 Qy 3781 CAGGCGCTGGGACCTGTAGAGCGAATGAAAAACAGCCCGACATCACTGGAGAAAGAC 3840
 Db 3602 CAGGCGCTGGGACCTGTAGAGCGAATGAAAAACAGCCCGACATCACTGGAGAAAGAC 3661
 Qy 3841 CACTCGCAAGCTCTCTGAGAGCAAGAACTGAAGAAAGTGAACCTGCAGCAT 3900
 Db 3662 CACTCGCAAGCTCTCTGAGAGCAAGAACTGAAGAAAGTGAACCTGCAGCAT 3721
 Qy 3901 GGCATCTCCAGTTTCTCTCTCTGAGCAACATACACAGGCTGGAGACTTTGGGGTTGTG 3960
 Db 3722 GGCATCTCCAGTTTCTCTCTCTGAGCAACATACACAGGCTGGAGACTTTGGGGTTGTG 3781
 Qy 3961 GTGCAGGCTTTCTTCCAAAAACACATCAGCCCACTGAGAGAGGGGACACTGACCTC 4020
 Db 3782 GTGCAGGCTTTCTTCCAAAAACACATCAGCCCACTGAGAGAGGGGACACTGACCTC 3841
 Qy 4021 ACCTTCTTCTCCAGAGCTTTTCACTTTCAGACCGCGGACATGCCAGATGCCATC 4080
 Db 3842 ACCTTCTTCTCCAGAGCTTTTCACTTTCAGACCGCGGACATGCCAGATGCCATC 3901
 Qy 4081 GCATCTGGAACAAAGCTGGAACCAAGTGGGCGGCAAGATCACTCCAACTTACACC 4140
 Db 3902 GCATCTGGAACAAAGCTGGAACCAAGTGGGCGGCAAGATCACTCCAACTTACACC 3961
 Qy 4141 CACAGCGGACCAAACTCAAGTGGGCTTCTCTGAGAGCCCTTCACTCAACCCCTCGG 4200
 Db 3962 CACAGCGGACCAAACTCAAGTGGGCTTCTCTGAGAGCCCTTCACTCAACCCCTCGG 4021
 Qy 4201 AACAGCGGATTTCTCCAGACCAAGCTGGAAGAGCCCCCGAGGTGCTTACTGGGCTGTC 4260
 Db 4022 AACAGCGGATTTCTCCAGACCAAGCTGGAAGAGCCCCCGAGGTGCTTACTGGGCTGTC 4081
 Qy 4261 CCAAGTGGCAGCGGAGTCCGCTTGTGTGGGCAATCCGAGAGTGTCTGAGAGTGC 4320
 Db 4082 CCAAGTGGCAGCGGAGTCCGCTTGTGTGGGCAATCCGAGAGTGTCTGAGAGTGC 4141
 Qy 4321 AGAGAGGTGGCCGAGAGGAAGATCTCAATGAGCTGCGCACTGGCCCTTAG 4377
 Db 4142 AGAGAGGTGGCCGAGAGGAAGATCTCAATGAGCTGCGCACTGGCCCTTAG 4198
 RESULT 5
 ABX97050
 ID ABX97050 standard; cDNA; 4425 BP.
 AC ABX97050;
 XX 20-MAY-2003 (first entry)
 XX
 DE Human NOV24c cDNA.
 XX
 KM NOVA; cytoskeletal; cardiant; antiarteriosclerotic; antiasthmatic; cancer;
 KM hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
 KM human; gene; ss.
 XX

OS Homo sapiens.
 XX
 PN W0200272757-A2.
 PD 19-SEP-2002.
 XX
 PF 08-MAR-2002; 2002W0-US06908.
 XX
 PR 08-MAR-2001; 2001US-274101P.
 PR 08-MAR-2001; 2001US-274194P.
 PR 08-MAR-2001; 2001US-274281P.
 PR 08-MAR-2001; 2001US-274322P.
 PR 09-MAR-2001; 2001US-274849P.
 PR 12-MAR-2001; 2001US-275235P.
 PR 13-MAR-2001; 2001US-275578P.
 PR 13-MAR-2001; 2001US-275579P.
 PR 13-MAR-2001; 2001US-275601P.
 PR 14-MAR-2001; 2001US-276000P.
 PR 16-MAR-2001; 2001US-276776P.
 PR 19-MAR-2001; 2001US-276994P.
 PR 20-MAR-2001; 2001US-277239P.
 PR 20-MAR-2001; 2001US-277321P.
 PR 20-MAR-2001; 2001US-277327P.
 PR 21-MAR-2001; 2001US-277791P.
 PR 22-MAR-2001; 2001US-277833P.
 PR 23-MAR-2001; 2001US-278152P.
 PR 26-MAR-2001; 2001US-278894P.
 PR 27-MAR-2001; 2001US-278999P.
 PR 27-MAR-2001; 2001US-279036P.
 PR 28-MAR-2001; 2001US-279344P.
 PR 30-MAR-2001; 2001US-279388P.
 PR 30-MAR-2001; 2001US-279995P.
 PR 30-MAR-2001; 2001US-280233P.
 PR 02-APR-2001; 2001US-280802P.
 PR 02-APR-2001; 2001US-280822P.
 PR 04-APR-2001; 2001US-280900P.
 PR 13-APR-2001; 2001US-283675P.
 PR 30-APR-2001; 2001US-287424P.
 PR 02-MAY-2001; 2001US-288066P.
 PR 03-MAY-2001; 2001US-288342P.
 PR 03-MAY-2001; 2001US-288528P.
 PR 15-MAY-2001; 2001US-291099P.
 PR 16-MAY-2001; 2001US-291240P.
 PR 16-MAY-2001; 2001US-291240P.
 PR 30-MAY-2001; 2001US-294485P.
 PR 31-MAY-2001; 2001US-294889P.
 PR 31-MAY-2001; 2001US-294899P.
 PR 18-JUN-2001; 2001US-299027P.
 PR 19-JUN-2001; 2001US-299303P.
 PR 19-JUN-2001; 2001US-299310P.
 PR 10-JUL-2001; 2001US-304354P.
 PR 31-JUL-2001; 2001US-309196P.
 PR 16-AUG-2001; 2001US-312903P.
 PR 10-SEP-2001; 2001US-318462P.
 PR 12-SEP-2001; 2001US-318770P.
 PR 27-SEP-2001; 2001US-325430P.
 PR 27-SEP-2001; 2001US-325681P.
 PR 18-OCT-2001; 2001US-330380P.
 PR 31-OCT-2001; 2001US-335301P.
 PR 14-NOV-2001; 2001US-332172P.
 PR 14-NOV-2001; 2001US-332271P.
 PR 14-NOV-2001; 2001US-332272P.
 PR 14-NOV-2001; 2001US-333184P.
 PR 14-NOV-2001; 2001US-333272P.
 PR 14-NOV-2001; 2001US-332094P.
 PR 03-DEC-2001; 2001US-337426P.
 PR 03-DEC-2001; 2001US-338092P.
 PR 04-DEC-2001; 2001US-337185P.
 PR 03-JAN-2002; 2002US-345705P.
 PR 07-MAR-2002; 2002US-0092900.
 XX
 XX (CURA-) CURAGEN CORP.

XX Padijari M, Spytek KA, Shenoy SG, Taupier RJ, Pena CEA, Li L,
PI Zethusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
PI Patturajan M, Gangoli E, Vernet CAM, Guo X, Tchernev V;
PI Fernandes ER, Caeman SJ, Malyankar UM, Gerlach V, Liu Y;
PI Anderson D, Spaderna SK, Calerton E, Burgess C, Lette M, Zhong H;
XX Alsobrook JP, Lepley DM, Rieger DK;
DR WPI; 2002-72332/78.
P-PSDB; ABUS5083.

XX NOVX polypeptides and polynucleotides, useful for preventing or
PT treating a disorder associated with aberrant NOVX expression or
PT activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or
PT bronchial asthma

XX Claim 13; Page 169-170; 1103pp; English.

XX This invention describes novel human NOVX polypeptides which have
CC cytotactic, cardiant, antiarteriosclerotic, antiasthmatic and
CC hypotensive activity. Pharmaceutical compositions comprising the NOVX
CC proteins or nucleic acid molecules or NOVX antibodies are useful for
CC preventing or treating a disorder associated with aberrant NOVX
CC expression or activity e.g., cancer, hypertension, atherosclerosis,
CC cardiomyopathy or bronchial asthma. The products of the invention can
CC be used for gene therapy or in a vaccine. ABX97008-ABX97185 are cDNA
CC fragments amplified and isolated by the PCR primers and probes
CC represented in ABX13460-ABX13462 and ABX97186-ABX97593. ABX97008-ABX97185
CC encode the NOVX proteins described in ABUS5041-ABUS5218.

XX Sequence 4425 BP; 1089 A; 1242 C; 1173 G; 921 T; 0 other;

XX Query Match 89.6%; Score 3923; DB 24; Length 4425;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 4181; Conservative 0; Mismatches 85; Indels 117; Gaps 9;

QY 1 ATGGGGCTGGCGGCGAGCATTTTCTCTGAGCTGCTGCTTCTGGGCGAAGGAGC 60
DB 16 ATGGGGCTGGCGGCGAGCATTTTCTCTGAGCTGCTGCTTCTGGGCGAAGGAGC 75
QY 61 CCTGAGATCCATCTCTCTTGAAGAAGTATGGAAGGCGAGCTATGGCGAGAGCC 120
DB 76 CCTGAGATCCATCTCTCTTGAAGAAGTATGGAAGGCGAGCTATGGCGAGAG-- 133
QY 121 CTGAAGAATCTCATTTCCATGCAACCCAAATTAATGAAGTGAATAGCTTCTTAA 180
DB 134 ----- 133
QY 181 TCAGTCACTCTCTGAAGCCTTCTGATATTAATTTGTGCGAGCATTTGCAATCTGAA 240
DB 134 -CAGTTCACCTCTCTGAAGCCTTCTGATATTAATTTGTGCGAGCATTTGCAATCTGAA 192
QY 241 ATTCTCCAGACCCGAGGAGCGGCGATCTGGAAGAAGCACTGGAAGAGGCGCAG 300
DB 193 ATGTGTCAGACCCGAGGAGCGGCGATCTGGAAGAAGCACTGGAAGAGGCGCAG 246
QY 301 CAGGTGTCATGGGAGTGAAGTCAATCTTTCAGACATCATCATATTTCACTGCTTCT 360
DB 247 CAGGTGTCATGGGAGTGAAGTCAATCTTTCAGACATCATCATATTTCACTGCTTCT 306
QY 361 GTTCCATATGCTGTGTCACATCTGGAAGAAGTCAATCCCAAGTGTGTCGAAGAC 420
DB 307 GTTCCATATGCTGTGTCACATCTGGAAGAAGTCAATCCCAAGTGTGTCGAAGAC 366
QY 421 TTGTGATTCAGGCTCAAGAACTGTGGAAGAAGTCAATGGAAGAAGCTGCAACTTT 480
DB 367 TTGTGATTCAGGCTCAAGAACTGTGGAAGAAGTCAATGGAAGAAGCTGCAACTTT 423
QY 481 CAATTTGACTGGAAGCTCATCATATGTCTTCTGATATGCAAGCCAGTGTACCTGTGC 540
DB 424 CAATTTGACTGGAAGCTCATCATATGTCTTCTGATATGCAAGCCAGTGTGTGTC 483
QY 541 CCTGTGCTCAAGAAATGGGCTTGGGCGGCGGTGATGAGCTGATGGGAGTGTGCTG 600

DB 484 CCTGTGCTCAAGAAATGGGCTTGGGCGGCGGTGATGAGCTGATGGGAGTGTGCTG 543
QY 601 GACTACTGACAGAGAGGTCCTCCAGAGCATTTGTAACTGTGTGACCTCTGTAGGTT 660
DB 544 GACTACTGACAGAGAGGTCCTCCAGAGCATTTGTAACTGTGTGACCTCTGTAGGTT 603
QY 661 GCAGAGTCTCTGTCATGATCAAGGCACTTGTGCTGAGCCCTGACAGAGCCCTGTAAT 720
DB 604 GCAGAGTCTCTGTCATGATCAAGGCACTTGTGCTGAGCCCTGACAGAGCCCTGTAAT 663
QY 721 TGTGAGAGAGACCAACCCGCTGGCCAGAGTGTGATGAGAGTGTCTTATCAGAAAGC 780
DB 664 TGTGAGAGAGACCAACCCGCTGGCCAGAGTGTGATGAGAGTGTCTTATCAGAAAGC 723
QY 781 TGAACAGCTCTGTCCTTCCAGACAGTACAGTACAGAGAGTCTTACCCGTGTTTC 840
DB 724 TGAACAGCTCTGTCCTTCCAGACAGTACAGTACAGAGAGTCTTACCCGTGTTTC 783
QY 841 CAGCTTTCTTATGAGACCAACCCATCTCTACCTGGAAGAGCCCGAGCTCAGAGAT 900
DB 784 CAGCTTTCTTATGAGACCAACCCATCT-----GACCCCGAGCTCAGAGAT 831
QY 901 TCTACACGCTGCGCTGCAATCTTGAATAGATGATGAGAGCAGAGAGAGAAAGAT 960
DB 832 TCTACACGCTGCGCTGCAATCTTGAATAGATGATGAGAGCAGAGAGAGAAAGAT 891
QY 961 GAGCCTTGAAGTATTAACACGGGAGGCCATGAATGTCCTCTCAGAGAGAGCCCTAT 1020
DB 892 GAGCCTTGAAGTATTAACACGGGAGGCCATGAATGTCCTCTCAGAGAGAGCCCTAT 951
QY 1021 CTGTGAGCTACAGAAACAGCACTACCTGACAGACTGAGAAACCCCAAGCAAGCTT 1080
DB 952 CTGTGAGCTACAGAAACAGCACTACCTGACAGACTGAGAAACCCCAAGCAAGCTT 1011
QY 1081 GAGTAAAGAGAGAGAGCGGAATCAGATGTCTGACAAAGACCCCTCCGATACGTTTCC 1140
DB 1012 GAGTAAAGAGAGAGAGCGGAATCAGATGTCTGACAAAGACCCCTCCGATACGTTTCC 1071
QY 1141 ACCTGAGTCTATGCTGAGAGCGGCTGACATCAAGTAAATGAGAGCCCTGAGTACTT 1200
DB 1072 ACCTGAGTCTATGCTGAGAGCGGCTGACATCAAGTAAATGAGAGAGCTGAGTACTT 1131
QY 1201 CTCAGGAGAGAGAGAGCGGCTGACATCAAGTAAATGAGAGAGCTGAGTACTT 1260
DB 1132 AGTACAGAGAGAGAGAGCGGCTGACATCAAGTAAATGAGAGAGCTGAGTACTT 1191
QY 1261 TACGAGAGAGAGAGAGCGGCTGACATCAAGTAAATGAGAGAGCTGAGTACTT 1320
DB 1192 TACGAGAGAGAGAGAGCGGCTGACATCAAGTAAATGAGAGAGCTGAGTACTT 1251
QY 1321 GCGAGATCTCTCGGGAATTAACCTTCTTGAAGGCTTCTGTGCTGCTGAGGAA 1380
DB 1252 GCGAGATCTCTCGGGAATTAACCTTCTTGAAGGCTTCTGTGCTGCTGAGGAA 1311
QY 1381 GAAACAGTCTTAATGCTCTTAAACCAAGCTGTGAGAGAGAGAGAGAGTGAATCTA 1440
DB 1312 GAAACAGTCTTAATGCTCTTAAACCAAGCTGTGAGAGAGAGAGAGAGTGAATCTA 1365
QY 1441 CTGTGAGAGAGAGAGAGCTGAGAGCTGAGAGATGAAGATGACAGAGATACCTTTCAG 1500
DB 1366 -----CAGGCGAGAGAGAGAGCTGAGAGATGAAGATGACAGAGATACCTTTCAG 1419
QY 1501 GAAAGTGAAGATTAATACCTGTTTATAGGCGGCAATGACCTGTGTGATTTCTGCAAT 1560
DB 1420 GAAAGTGAAGATTAATACCTGTTTATAGGCGGCAATGACCTGTGTGATTTCTGCAAT 1479
QY 1561 GATCTGTGACATATCTCTCCGAGACTTCAAGAGAACTTGAAGAAAGGCGCTGAGATC 1620
DB 1480 GATCTGTGACATATCTCTCCGAGACTTCAAGAGAACTTGAAGAAAGGCGCTGAGATC 1539
QY 1621 CTCATGCTGAGAGTCTCTGAGGACTTGTGAACCTGTGATGAGTGTGATGATGATGAT 1680

Db 1540 CTCATCTGAGGTTCTCGGGCAATTTGTGAACCTGGTGAACGGTCTTGAGATGTCAC 1599
Qy 1681 CTGAGGAGCTGTACAGAGAAAAAGTCTACTGCCCAAGGATGATCCCAAGTCTCTG 1740
Db 1600 CTGAGGAGCTGTACAGAGAAAAAGTCTACTGCCCAAGGATGATCCCAAGTCTCTG 1659
Qy 1741 TGTCCCTGTCTCTGAAGTTGATGATTAACCAAGAACTGTCTACCTCTCATGAAATTC 1800
Db 1660 TGTCCCTGTCTCTGAAGTTGATGATTAACCAAGAACTGTCTACCTCTCATGAAATTC 1719
Qy 1801 AACAGAAGTTTACAGAGAAAGCCCACTGATTAAGATGGGCAATATGACAAGG 1860
Db 1720 AACAGAAGTTTACAGAGAAAGCCCACTGATTAAGATGGGCAATATGACAAGG 1779
Qy 1861 GAAGATTTTACTGTGTTGTGACAGCCGTTCTTGAACAGTGAACAGCCAAAGACTCG 1920
Db 1780 GAAGATTTTACTGTGTTGTGACAGCCGTTCTTGAACAGTGAACAGCCAAAGACTCG 1839
Qy 1921 GAAGATTTGCTGCAACTCTTTCTGCTCTGACTGTTTCCACTTCAGCAGAGTCT 1980
Db 1840 GAAGATTTGCTGCAACTCTTTCTGCTCTGACTGTTTCCACTTCAGCAGAGTCT 1899
Qy 1981 CACTCCGAGCAGCAGTGTCTCTGGAACAATATGCTGAGCCTGTTGGCAGAGACG 2040
Db 1900 CACTCCGAGCAGCAGTGTCTCTGGAACAATATGCTGAGCCTGTTGGCAGAGACG 1959
Qy 2041 ACTGCTATAGTTTGAACAGATCAATATCAATGCTGGAACAGAGTCCA---GCG 2097
Db 1960 ACTGCTATAGTTTGAACAGATCAATATCAATGCTGGAACAGAGTGAAGTGGCG 2019
Qy 2098 TTTCTGAGACCTTACAGAAACAGATCAGAGTCTGAGGAGCTGGCTGCAATGAGGAG 2157
Db 2020 TTTCTGAGACCTTACAGAAACAGATCAGAGTCTGAGGAGCTGGCTGCAATGAGGAG 2079
Qy 2158 AGAGCCCTTCTGCTTGAACCTTACCTGATGATGCTGAGACCTGCAAGATCCAA 2217
Db 2080 AGAGCCCTTCTGCTTGAACCTTACCTGATGATGCTGAGACCTGCAAGATCCAA 2139
Qy 2218 GTTGTGCTCTCTGAGGAGATTCCTGACCGCTGGCAATGGAATTTGGCTCAACCGAG 2277
Db 2140 GTTGTGCTCTCTGAGGAGATTCCTGACCGCTGGCAATGGAATTTGGCTCAACCGAG 2199
Qy 2278 GACTCCCGCATGTCAACACAGATACGAGGAGCTGTATACAGTGAAGAGGAGGAGCG 2337
Db 2200 GACTCCCGCATGTCAACACAGATACGAGGAGCTGTATACAGTGAAGAGGAGGAGCG 2259
Qy 2338 TCCCTGAGATGTGACCACTTAATTAATTCCTGAGGAGTTTAAACAGAACTGACA 2397
Db 2260 TCCCTGAGATGTGACCACTTAATTAATTCCTGAGGAGTTTAAACAGAACTGACA 2319
Qy 2398 GGCCTACGCGTGGGCAAGGATGCAATGACAGAACTCAATTCCTCATATCAAGCTGT 2457
Db 2320 GGCCTACGCGTGGGCAAGGATGCAATGACAGAACTCAATTCCTCATATCAAGCTGT 2379
Qy 2458 CCGGAGCAAAAGCTGAGATCTTATGAGCCAACTGCAAACTGATGAGAGATGAAA 2517
Db 2380 CCGGAGCAAAAGCTGAGATCTTATGAGCCAACTGCAAACTGATGAGAGATGAAA 2439
Qy 2518 GATGATCATAGATTAATTTTCATGAAAGCTGAAAGTCAATCAAGTGTGATGAGGC 2577
Db 2440 GATGATCATAGATTAATTTTCATGAAAGCTGAAAGTCAATCAAGTGTGATGAGGC 2499
Qy 2578 AGGATTTATGATCTTCTGCAAGATTCGAATGTGATTTCTGAGGCAACTTTGTGAC 2637
Db 2500 AGGATTTATGATCTTCTGCAAGATTCGAATGTGATTTCTGAGGCAACTTTGTGAC 2559
Qy 2638 CATCTCCGCAATGCTTGAAGCTCTGATAGAGAGGTCGAGAGCTGTGATCAACTC 2697
Db 2560 CATCTCCGCAATGCTTGAAGCTCTGATAGAGAGGTCGAGAGCTGTGATCAACTC 2619
Qy 2698 GTGAGCTTCTTGAACCCCACTATCATGCGAGAGTGTCTCTGAGAAACCAAGACAGTGC 2757
Db 2620 GTGAGCTTCTTGAACCCCACTATCATGCGAGAGTGTCTCTGAGAAACCAAGACAGTGC 2679

Qy 2758 CAAATGAGAGGCGACGCTTTTGTGAACTGCGTTCTGACCTGCGGAGAACTCCCA 2817
Db 2680 CAAATGAGAGGCGACGCTTTTGTGAACTGCGTTCTGACCTGCGGAGAACTCCCA 2739
Qy 2818 GAATGAGGAGGCTGAGGCTTCAAGCCGAGCTTACCGAGCAGCATGCGGAGCTGGT 2877
Db 2740 GAATGAGGAGGCTGAGGCTTCAAGCCGAGCTTACCGAGCAGCATGCGGAGCTGGT 2799
Qy 2878 GGGTCAAGGCGCTATGACAGGAGAGGATTTCTGTGTGTCTGTGAGCCCTTCTCCAG 2937
Db 2800 GGGTCAAGGCGCTATGACAGGAGAGGATTTCTGTGTGTGTGTGAGCCCTTCTCCAG 2859
Qy 2938 AACATCAGCTCCCTGCTGCGGAGTGGGCTCCAGATGAGTCTCTTTGCCCCAGAC 2997
Db 2860 AACATCAGCTCCCTGCTGCGGAGTGGGCTCCAGATGAGTCTCTTTGCCCCAGAC 2919
Qy 2998 TGCATCCACCCCAATCAGAAATTCATCCAGCTGGCCAGAGCCCTTTGGAACAAATAG 3057
Db 2920 TGCATCCACCCCAATCAGAAATTCATCCAGCTGGCCAGAGCCCTTTGGAACAAATAG 2979
Qy 3058 CTGGAACCACTTGGAAACCAAGAGCCCTGAGCTTGAAGCAGAGATGCCCATAC 3117
Db 2980 CTGGAACCACTTGGAAACCAAGAGCCCTGAGCTTGAAGCAGAGATGCCCATAC 3039
Qy 3118 TGTCCCACTCAGAAATGAGCCCTTCTGAGAAACCCCTGCGAATGTAACTACAGTACCC 3177
Db 3040 TGTCCCACTCAGAAATGAGCCCTTCTGAGAAACCCCTGCGAATGTAACTACAGTACCC 3099
Qy 3178 ATCAAGCCAGCCATTTGAGAACTGGGCACTGATCTTCTGTATCAGAGTGAAGCTTC 3237
Db 3100 ATCAAGCCAGCCATTTGAGAACTGGGCACTGATCTTCTGTATCAGAGTGAAGCTTC 3159
Qy 3238 AATAGTGTTCACCTCTGTCCACAGCTCCGACAGCAGACATTCAAAGTGTGGCCGCC 3297
Db 3160 AATAGTGTTCACCTCTGTCCACAGCTCCGACAGCAGACATTCAAAGTGTGGCCGCC 3219
Qy 3298 CTGAGTGACTCTGTACTACAGAGTGGGAGCTGACCAAAACAATCCAGATGACTTACC 3357
Db 3220 CTGAGTGACTCTGTACTACAGAGTGGGAGCTGACCAAAACAATCCAGATGACTTACC 3279
Qy 3358 ACATCTTGAAGGAGACTCTTTGAGACATTTGAGGAGATGAGAACTTGAAGTCAACC 3417
Db 3280 ACATCTTGAAGGAGACTCTTTGAGACATTTGAGGAGATGAGAACTTGAAGTCAACC 3339
Qy 3418 ACATCTGCAACATTTGAGAAAGTTCAACCTTACCTCTTGGCTTCTTACAGACACC 3477
Db 3340 ACATCTGCAACATTTGAGAAAGTTCAACCTTACCTCTTGGCTTCTTACAGACACC 3399
Qy 3478 TGGAGGAGGAGCAGCAGAGCTAAATGTGGCAGCGAGAGGAGCCAGAGCTAGGAGATGCCA 3537
Db 3400 TGGAGGAGGAGCAGCAGAGCTAAATGTGGCAGCGAGAGGAGCCAGAGCTAGGAGATGCCA 3459
Qy 3538 GCCAGAGCCCTGGAGCTGTGAGCGAATGAAACAGGCCCC--GACATCAACTGGAG 3594
Db 3460 GCCAGAGCCCTGGAGCTGTGAGCGAATGAAACAGGCCCCGAGAGCATCAACTGGAG 3519
Qy 3595 AAAGATGGAAGTGTGTACACTTCTTATTTGGGGTCAAGCATTTGTGTCTTACTGTGAG 3654
Db 3520 AAAGATGGAAGTGTGTACACTTCTTATTTGGGGTCAAGCATTTGTGTCTTACTGTGAG 3579
Qy 3655 AATCGAGGAGCCCACTTGGCAGCAGATATGTTTCAACATCCCAAGGAGCCCTGAGCATC 3714
Db 3580 AATCGAGTGGG-----GAATATGTTTCAACATCCCAAGGAGCCCTGAGCATC 3627
Qy 3715 CTCTTGAGAGGCTCCCAAGAGGCTTTCGTCAAGTGTGAGAGTCAATGAGTGTGAGTGC 3774
Db 3628 CTCTTGAGAGGCTCCCAAGAGGCTTTCGTCAAGTGTGAGAGTCAATGAGTGTGAGTGC 3687
Qy 3775 CTGTACAGGAGCAGAGGCGGAAATGTGCATGCTGTGACATGCAACATGCACTTGC 3834
Db 3688 CTGTACAGGAGCAGAGGCGGAAATGTGCATGCTGTGACATGCAACATGCACTTGC 3747


```

QY 3835 CTGAGACACTCGCAAAAGCTCTCTGAGAGCAAGAACTGAAAGAACTGAACTGAACTG 3894
DB 3748 CTGAGACACTCGCAAAAGCTCTCTGAGAGCAAGAACTGAAAGAACTGAACTGAACTG 3807
QY 3895 CAGATGGCATCTCCAGTTTCTCTACTGAGCAACAAATACACAGCGCTGAGCACTTGGC 3954
DB 3808 CAGATGGCATCTCCAGTTTCTCTACTGAGCAACAAATACACAGCGCTGAGCACTTGGC 3867
QY 3955 GTTGGTGGAGCGCTTCTCTCAAAAACACACTCACTGAGAGAGAGAGAGAGAGAGAG 4014
DB 3868 GAGACTCGCAAAAGCTCTCTCTCAAAAACACACTCACTGAGAGAGAGAGAGAGAGAG 3924
QY 4015 GACCTCACTCTCTCTCTCTGAGAGCTGTTTCACTTCTCAGACCGCGGAGATGCGAGATG 4074
DB 3925 GACCTCACTCTCTCTCTCTGAGAGCTGTTTCACTTCTCAGACCGCGGAGATGCGAGATG 3984
QY 4075 GGCATGGCATCTGGAACAAATGCTGGAACCAAGTGGCGGCAAGACTCTCCAAACAC 4134
DB 3985 GGCATGGCATCTGGAACAAATGCTGGAACCAAGTGGCGGCAAGACTCTCCAAACAC 4044
QY 4135 TTGACCCAGAGCGGAGCAAACTCAAGTGGCGGCTCTGAGAGAGAGAGAGAGAGAG 4194
DB 4045 TTGACCCAGAGCGGAGCAAACTCAAGTGGCGGCTCTGAGAGAGAGAGAGAGAGAG 4104
QY 4195 CTGCGGAACAGCGGATTTCTCTCCAGACCAAGGCTGGAAGAGAGAGAGAGAGAGAG 4254
DB 4105 CTGCGGAACAGCGGATTTCTCTCCAGACCAAGGCTGGAAGAGAGAGAGAGAGAGAG 4164
QY 4255 GCTGTCTCCAGTGGAGCGGAGAGTGGCGCTTGTGTGGTGGCATTCATCGGAGACAGTGTGG 4314
DB 4165 GCTGTCTCCAGTGGAGCGGAGAGTGGCGCTTGTGTGGTGGCATTCATCGGAGACAGTGTGG 4224
QY 4315 AGGTGACGAGAGAGTGGCGGAGAGAGATCTCCAAATGAGAGAGAGAGAGAGAGAG 4374
DB 4225 AGGTGACGAGAGAGTGGCGGAGAGAGATCTCCAAATGAGAGAGAGAGAGAGAGAG 4284
QY 4375 TAG 4377
DB 4285 TAG 4287

RESULT 6
AAD37410
ID AAD37410 standard; cDNA; 3648 BP.
XX
AC AAD37410;
XX
DT 27-AUG-2002 (first entry)
XX
DE Human phospholipase-like enzyme encoding cDNA.
XX
KW Human; phospholipase-like enzyme; cancer; inflammation; Pick's disease;
KW cardiovascular disorder; central nervous system disorder; brain injury;
KW chronic obstructive pulmonary disease; cerebrovascular disease; dementia;
KW Alzheimer's disease; Parkinson's disease; corticobasal degeneration;
KW motor neuron disease; Huntington's disease; Creutzfeldt Jacob dementia;
KW schizophrenia; Korsakoff's psychosis; pain; epilepsy; multiple sclerosis;
KW sclerotic; stroke; age associated memory impairment; allergy; asthma;
KW allergic rhinitis; hay fever; atopic dermatitis; cardiovascular disease;
KW anaphylaxis; inflammation; acute respiratory distress syndrome; diabetes;
KW chronic obstructive pulmonary disease; emphysema; obesity; anorexia;
KW overweight; cachexia; bulimia; hypertension; coronary artery disease;
KW type-II diabetes; hyperlipidaemia; gall bladder disease; osteoarthritis;
KW gout; sleep apnoea; respiratory problem; polycystic ovarian syndrome;
KW chromolytic disease; reduced fertility; pregnancy; stress incontinence;
KW hirsutism; menstrual irregularity; depression; enzyme; gene; ss.
XX
OS Homo sapiens.
XX
XX Location/Qualifiers
XX Key 1..3648
XX CDS
XX FT /product= "Human phospholipase-like enzyme"
XX FT

```

```

FT /note= "CDS does not include start and stop codon"
FT /partial
XX
XX NO200231161-A2.
XX
XX 18-APR-2002.
XX
XX 09-OCT-2001, 2001WO-BP11641.
XX
XX 10-OCT-2000, 2000US-238445P.
XX 26-DEC-2000, 2000US-257293P.
XX
XX (FARB ) BAYER AG.
XX
XX Zhu Z.
XX
XX WPI, 2002-426287/45.
XX P-PSDB; AAB22860.
XX
XX New human phospholipase-like enzyme polypeptide useful for screening
XX agents, and in the treatment of cancer, inflammation, diabetes,
XX obesity, a central nervous system disorder, or a cardiovascular
XX disorder
XX
XX Claim 19, Fig 1, 144p; English.
XX
XX The present invention relates to novel human phospholipase-like enzymes
XX and polynucleotides encoding such proteins. Sequences of the invention
XX are useful for producing a medicament for modulating the activity of
XX phospholipase in a disease such as cancer, inflammation, cardiovascular
XX disorders, chronic obstructive pulmonary diseases, central nervous system
XX (CNS) disorders such as brain injuries, cerebrovascular disease, dementia
XX (Alzheimer's disease), Parkinson's disease, corticobasal degeneration,
XX motor neuron disease, Pick's disease, Huntington's disease, Creutzfeldt
XX Jacob dementia, schizophrenia with dementia, Korsakoff's psychosis,
XX pain associated with CNS (e.g. epilepsy, failed back surgery syndrome,
XX sclerotic), multiple sclerosis, stroke, age associated memory impairment,
XX allergic disease including asthma, allergic rhinitis (hay fever), atopic
XX dermatitis, anaphylaxis and inflammation, cardiovascular disease, chronic
XX obstructive pulmonary disease, acute respiratory distress syndrome, gout,
XX diabetes, emphysema or obesity. They are also used for treating anorexia,
XX overweight, cachexia, bulimia, hypertension, type-II diabetes, coronary
XX artery disease, hyperlipidaemia, gall bladder disease, osteoarthritis,
XX sleep apnoea and respiratory problems, cancer (e.g. breast, prostate,
XX colon cancer), thrombolytic disease, reduced fertility, polycystic
XX ovarian syndrome, complications of pregnancy, menstrual irregularity,
XX hirsutism, stress incontinence and depression. The present sequence is
XX human phospholipase like enzyme encoding cDNA.
XX
XX Sequence 3648 BP, 899 A, 1032 C, 947 G, 770 T, 0 other;
XX
XX Query Match 79.1%; Score 3461.2; DB 24; Length 3648;
XX Best Local Similarity 97.5%; Pred. No. 0;
XX Matches 3558; Conservative 0; Mismatches 33; Indels 57; Gaps 2;
XX
XX 775 GAAGCTGGAACAGCTCTCTGAGAGCAAGTAAAGAGAGAGAGAGTCTTACCGTG 834
XX 1 GAAGCTGGAACAGCTCTCTGAGAGCAAGTAAAGAGAGAGAGAGTCTTACCGTG 60
XX
XX 835 GTTTCAGAGCTTCTCTATGAGACACCGCATCTCTCACTCGGAGAGAGAGAGAGAG 894
XX 61 GTTTCAGAGCTTCTCTATGAGACACCGCATCTCTCACTCGGAGAGAGAGAGAGAG 120
XX
XX 895 CAGATTTACACAGCTGAGCATCTCTGGAATAGATGATGAGAGAGAGAGAGAG 954
XX 121 CAGATTTACACAGCTGAGCATCTCTGGAATAGATGATGAGAGAGAGAGAGAG 180
XX
XX 955 AAAGATGAGCCATTTAGTGAACACCGGAGAGAGAGAGAGAGAGAGAGAGAG 1014
XX 181 AAAGATGAGCCATTTAGTGAACACCGGAGAGAGAGAGAGAGAGAGAGAGAG 240
XX
XX 1015 CCTATCTGTTGAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1074
XX

```

Db 241 CCTATCTGTACGTA CAGAA CAGCACTGACGACGTA CAGAA CCCCAGAC 300
Qy 1075 AGCTTG-----AGTAAGAA 1092
Db 301 AAGCTTGAAGAAATTCATTTCTTATTTCCGTGTGTCTCATTTCAAGTAAGAA 360
Qy 1093 GAGACGAAATCAGATGCTGACAAAGACCCCTCGATACGATGCTCCAGCTCAT 1152
Db 361 GAGACGAAATCAGATGCTGACAAAGACCCCTCGATACGATGCTCCAGCTCAT 420
Qy 1153 AGGCTGAAGCCGCTGACATCAAGATGTAATGAGCCCTGGTGAATCTTCA CCGCAGC 1212
Db 421 AGGCTGAAGCCGCTGACATCAAGATGTAATGAGCCCTGGTGAATCTTCA CCGCAGC 480
Qy 1213 AATGAGGCGGGGTCAACCTCGGAGAGTCTTGAACGCTTGAATCACTGATCCGAGGCTG 1272
Db 481 AATGAGGCGGGGTCAACCTCGGAGAGTCTTGAACGCTTGAATCACTGATCCGAGGCTG 540
Qy 1273 TCCTGAGCGTCCGCGAGATGAGACATCGCACCCGTTACCACTGCGGACATCTC 1332
Db 541 TCCTGAGCGTCCGCGAGATGAGACATCGCACCCGTTACCACTGCGGACATCTC 600
Qy 1333 CGGGAATTCAAACCTTCCCTGAAGGCTTCTGTGTGCACTGGGAAAGAA CCAATCT 1392
Db 601 CGGGAATTCAAACCTTCCCTGAAGGCTTCTGTGTGCACTGGGAAAGAA CCAATCT 660
Qy 1393 AATGCTCTTAAACCAAGGCTGTGGCAGAGCCGAGCTGAGATCACTGCTCCAGGCT 1452
Db 661 AATGCTCTTAAACCAAGGCTGTGGCAGAGCCGAGCTGAGATCACTGCTCCAGGCT 720
Qy 1453 AGGAGCTGTGTGACCTGATGAGATGACACAGAGATACCTTTCAGAAAGCTGAG 1512
Db 721 AGGAGCTGTGTGACCTGATGAGATGACACAGAGATACCTTTCAGAAAGCTGAG 780
Qy 1513 ATAATAACCTGTTTATAGGCGCAATGACTCTGTGATTTCTGCAATGATCTGTCCAC 1572
Db 781 ATAATAACCTGTTTATAGGCGCAATGACTCTGTGATTTCTGCAATGATCTGTCCAC 840
Qy 1573 TATTCCTCCAGAACTTCA CAGACCAATGGAAGGCGCTGACATCTCCAGTGTGAG 1632
Db 841 TATTCCTCCAGAACTTCA CAGACCAATGGAAGGCGCTGACATCTCCAGTGTGAG 900
Qy 1633 GTTCTCGGGCATTTGTGAACCTGTGACGATGCTTGAATCTTCAACTGAGGAGCTG 1692
Db 901 GTTCTCGGGCATTTGTGAACCTGTGACGATGCTTGAATCTTCAACTGAGGAGCTG 960
Qy 1693 TACCAAGAGAAAGATCTACCTGCCAAGATGATCTCAAGTCTGTGTCTGTGTCT 1752
Db 961 TACCAAGAGAAAGATCTACCTGCCAAGATGATCTCAAGTCTGTGTCTGTGTCT 1020
Qy 1753 CTGAAGTTGATGATTA CTA CAGAACTTGCTACCTCATCGAATTCACAAAGATTT 1812
Db 1021 CTGAAGTTGATGATTA CTA CAGAACTTGCTACCTCATCGAATTCACAAAGATTT 1080
Qy 1813 CAGAGAGAGACCCCACTGATGAGATGAGGCGATATGACAAAGGAAATTTACT 1872
Db 1081 CAGAGAGAGACCCCACTGATGAGATGAGGCGATATGACAAAGGAAATTTACT 1140
Qy 1873 GTGCTGTGACGCGCTTCTTGAAGCGTGAACATGCCAAAGACTTGGAGATGCT 1932
Db 1141 GTGCTGTGACGCGCTTCTTGAAGCGTGAACATGCCAAAGACTTGGAGATGCT 1200
Qy 1933 GACAACCTTTCTTCCCTGCTGATGTTTCCATTCAGCAGAGTCTCACTCCGAGCA 1992
Db 1201 GACAACCTTTCTTCCCTGCTGATGTTTCCATTCAGCAGAGTCTCACTCCGAGCA 1260
Qy 1993 GCGAGTCTCTCTGAA CAAATATCTGAGGCTGTTGCGCAGAGAGCAACTGCTATAG 2052
Db 1261 GCGAGTCTCTCTGAA CAAATATCTGAGGCTGTTGCGCAGAGAGCAACTGCTATAG 1320
Qy 2053 TTTGAAACAGATCAATATCACTGTCGAA CCAAGGTCGACCGCTTCTGAGGACCTAC 2112
Db 1321 TTTGAAACAGATCAATATCACTGTCGAA CCAAGTCTTCTGTCCCTTCAACCTG 1380

Qy 2113 AAGA CAG-----CATGAGGCTCATGAGGACCTGGCTGCCATGACAGGAC 2157
Db 1381 AGATCTGGAGATTTGTTCTATGCTCCCTTAGGGGTCAAGGACCTGCTCCATGACAGGAC 1440
Qy 2158 AGAGCCCTTCTGCTTGA CACCTTACCTCAAGTCAATGCTGAGACCTGAGACATTCGA 2217
Db 1441 AGAGCCCTTCTGCTTGA CACCTTACCTCAAGTCAATGCTGAGACCTGAGACATTCGA 1500
Qy 2218 GTTGTGCTGCTTGGGGGATTTCTGACCGCTGGCAATGGAATTTGGCTCCAA CCAAGAC 2277
Db 1501 GTTGTGCTGCTTGGGGGATTTCTGACCGCTGGCAATGGAATTTGGCTCCAA CCAAGAC 1560
Qy 2278 GACCTCCCGATGTCAC CACAGATATCGGGGACTGTATACAGTCAAGGAGGGAGCGC 2337
Db 1561 GACCTCCCGATGTCAC CACAGATATCGGGGACTGTATACAGTCAAGGAGGGAGCGC 1620
Qy 2338 TCCTGAGAAATGAC CACCTTACCTAATATCTTCCGAGATTTA CAGAA CTTGACA 2397
Db 1621 TCCTGAGAAATGAC CACCTTACCTAATATCTTCCGAGATTTA CAGAA CTTGACA 1680
Qy 2398 GGTACGCGCTGGGACAGGATGATGCCAATGACACGAATGATTTCTCATCAAGCTGT 2457
Db 1681 GGTACGCGCTGGGACAGGATGATGCCAATGACACGAATGATTTCTCATCAAGCTGT 1740
Qy 2458 CCGGAGCAAAAGCTGAGATCTTATGAGCCAACTGCAATGCAAGATGAAA 2517
Db 1741 CCGGAGCAAAAGCTGAGATCTTATGAGCCAACTGCAATGCAAGATGAAA 1800
Qy 2518 GATGATCATAGATAATTTCAATGAAGCTGAGAGCTCATCAAGTGTGATCGAGGC 2577
Db 1801 GATGATCATAGATAATTTCAATGAAGCTGAGAGCTCATCAAGTGTGATCGAGGC 1860
Qy 2578 AGCGATTTATGATCACTGACAGAGTGTGAATCTGTATTTCTGACGCCAACTTTGTGAC 2637
Db 1861 AGCGATTTATGATCACTGACAGAGTGTGAATCTGTATTTCTGACGCCAACTTTGTGAC 1920
Qy 2638 CATCTCCGCAATGCTTGAAGCTGTCGATGAGAGAGTGTCCAGAGCTCGTCAACCTC 2697
Db 1921 CATCTCCGCAATGCTTGAAGCTGTCGATGAGAGAGTGTCCAGAGCTCGTCAACCTC 1980
Qy 2698 GTGACCTTCTGAACCCCACTATCATGCGGACAGTGTCTGTGGAAACCAGCAAGTGC 2757
Db 1981 GTGACCTTCTGAACCCCACTATCATGCGGACAGTGTCTGTGGAAACCAGCAAGTGC 2040
Qy 2758 CCAATGACAGGCGCAGCTTTGTGTAACTGCTGTGACCTCTGGGGGAACTCCCA 2817
Db 2041 CCAATGACAGGCGCAGCTTTGTGTAACTGCTGTGACCTCTGGGGGAACTCCCA 2100
Qy 2818 GAGCTAGCAGGCTGAGGCTTCA GCGAGCTTACCGAGAGCAATGCGGAGCTGTG 2877
Db 2101 GAGCTAGCAGGCTGAGGCTTCA GCGAGCTTACCGAGAGCAATGCGGAGCTGTG 2160
Qy 2878 GGGTCAAGGCGCTATGACAGCAGAGAGACTTCTGTGTGTGTGACGCCCTTCTCAG 2937
Db 2161 GGGTCAAGGCGCTATGACAGCAGAGAGACTTCTGTGTGTGTGACGCCCTTCTCAG 2220
Qy 2938 AACATCAGCTCCCTGTCTGTGGGATGGCTCCAGATACGTCCTTCTTGGCCAGAC 2997
Db 2221 AACATCAGCTCCCTGTCTGTGGGATGGCTCCAGATACGTCCTTCTTGGCCAGAC 2280
Qy 2998 TGCATCCACCAATCAGAAATTCATCCAGCTGAGGCGAGGCGCTTGGACCAATATG 3057
Db 2281 TGCATCCACCAATCAGAAATTCATCCAGCTGAGGCGAGGCGCTTGGACCAATATG 2340
Qy 3058 CTTGAACCACTTGAAGCAAA CAGAGACCTTGA CTTGAGACAGAGATGCCATCAC 3117
Db 2341 CTTGAACCACTTGAAGCAAA CAGAGACCTTGA CTTGAGACAGAGATGCCATCAC 2400
Qy 3118 TGTCCACTCAAGATGAGCTTCTGAGAACCTTGGGAATGTAACTACAGTACCC 3177
Db 2401 TGTCCACTCAAGATGAGCTTCTGAGAACCTTGGGAATGTAACTACAGTACCC 2460

QY 3178 ATCAAGCCAGCATTTAGAACTGAGGAGTGAATCTCTGTGTAGAGAGTGAAGGCTTCC 3237
 Db 2461 ATCAAGCCAGCATTTAGAACTGAGGAGTGAATCTCTGTGTAGAGAGTGAAGGCTTCC 2520
 QY 3238 AATTGTGTTCCAACTCTGTCTACCAAGTCCGACGACGACATCAAAAGTGTGGCGCC 3297
 Db 2521 AATTGTGTTCCAACTCTGTCTACCAAGTCCGACGACGACATCAAAAGTGTGGCGCC 2580
 QY 3298 CTGGGACTCTCTGACTACAGAGTGGAGCTGACCAACCACTCCAGTACCTACCC 3357
 Db 2581 CTGGGACTCTCTGACTGTAGTATGGAGCTGACCAACCACTCCAGTACCTACCC 2640
 QY 3358 ACATCTTGAAGGGAGCTCTTGTGAGCATTTGAGGGATGGGAATTGAGACTCACACC 3417
 Db 2641 ACATCTTGAAGGGAGCTCTTGTGAGCATTTGAGGGATGGGAATTGAGACTCACACC 2700
 QY 3418 ACATCTGCCCCAATCTTGAAGAAAGTTCAACCTTACTCTTGGCTTCTCTACCAAGACC 3477
 Db 2701 ACATCTGCCCCAATCTTGAAGAAAGTTCAACCTTACTCTTGGCTTCTCTACCAAGACC 2760
 QY 3478 TGGGAGGGGACAGCAGGACTAAATGTGGCAGCGGAGGGGCGAGAGCTAGGGAGCATGCCA 3537
 Db 2761 TGGGAGGGGACAGCAGGACTAAATGTGGCAGCGGAGGGGCGAGAGCTAGGGAGCATGCCA 2820
 QY 3538 GCCCAGGCTGGGACCTGTGTAGACGGAATGAAAAACAGCCCGACATCAACCTGAGAAA 3597
 Db 2821 GCCCAGGCTGGGACCTGTGTAGACGGAATGAAAAACAGCCCGACATCAACCTGAGAAA 2880
 QY 3598 GACTGGAAGCTGTCTCAACTTTTATGGGGTCAACGACTTGTCTATTACTGTAGAAAT 3657
 Db 2881 GACTGGAAGCTGTCTCAACTTTTATGGGGTCAACGACTTGTCTATTACTGTAGAAAT 2940
 QY 3658 CCGGAGGCCCACTGGGCGCAGGAAATATGTTTCAAGACATCCAAAGGCCCTGAGACATGCTC 3717
 Db 2941 CCGGAGGCCCACTGGGCGCAGGAAATATGTTTCAAGACATCCAAAGGCCCTGAGACATGCTC 3000
 QY 3718 TCTGAGAGAGTCCCAAGGGCTTTTGTCTCAACGTGTGAGAGCTGAGAGTGGTACGCTG 3777
 Db 3001 TCTGAGAGAGTCCCAAGGGCTTTTGTCTCAACGTGTGAGAGCTGAGAGTGGTACGCTG 3060
 QY 3778 TACCAGGGCCAAAGCGGGAAATGTGCATGTGGCAGCTCAGAACATCTGACTTGCCTC 3837
 Db 3061 TACCAGGGCCAAAGCGGGAAATGTGCATGTGGCAGCTCAGAACATCTGACTTGCCTC 3120
 QY 3838 AGACACTTCGAAAGCTCCCTGTGAGAGACGAAAGACTGAAAGAAAGTGAACCTTCAG 3897
 Db 3121 AGACACTTCGAAAGCTCCCTGTGAGAGACGAAAGACTGAAAGAAAGTGAACCTTCAG 3180
 QY 3898 CATGGCATCTCCAGTTTCTCTACTGAGCAACCAATACACAGCGGTGAGGACTTGGCGGTT 3957
 Db 3181 CATGGCATCTCCAGTTTCTCTACTGAGCAACCAATACACAGCGGTGAGGACTTGGCGGTT 3240
 QY 3958 GTGGTGAAGCTTTCTTCAAAAACACATCAACCCACTGAAAGAGAGGGGACACTGAC 4017
 Db 3241 GTGGTGAAGCTTTCTTCAAAAACACATCAACCCACTGAAAGAGAGGGGACACTGAC 3300
 QY 4018 CTCACTTTCTTCTCCGAGAGCTTTTCACTTCAACAGCGGGGACTGCCAGATGGCC 4077
 Db 3301 CTCACTTTCTTCTCCGAGAGCTTTTCACTTCAACAGCGGGGACTGCCAGATGGCC 3360
 QY 4078 ATGCACTCTGGAACAACATGCTGGAACAAGTGGGCGGCAAGACTTACTCCAACTTC 4137
 Db 3361 ATGCACTCTGGAACAACATGCTGGAACAAGTGGGCGGCAAGACTTACTCCAACTTC 3420
 QY 4138 ACCCAGCGGAGCAAACTCAAGTGCCTCTCTCTGAGAGCCCTTACTCTACACCTTG 4197
 Db 3421 ACCCAGCGGAGCAAACTCAAGTGCCTCTCTCTGAGAGCCCTTACTCTACACCTTG 3480
 QY 4198 CGGAAACGCGGATGCTCTCCAGACCAAGCTGGAAGAGCCCCCGAGGTCTCTACTGGGCT 4257
 Db 3481 CGGAAACGCGGATGCTCTCCAGACCAAGCTGGAAGAGCCCCCGAGGTCTCTACTGGGCT 3540
 QY 4258 GTCCCACTGGCAGGAGGTGCGGCTTGTGTGGGCAATTCGCGGACAGTGTCTGGAGG 4317

Db 3541 GTCCAGTGGCAGGCGGAGTTCGGCTTGTGTGGGCAATTCAGGAGAGTGTCTGGAGG 3600
 QY 4318 TGCAGAGAGAGTGGCCCGGAGGAAAGATCTTCATGAGCTGCGCACT 4365
 Db 3601 TGCAGAGAGAGTGGCCCGGAGGAAAGATCTTCATGAGAGCTGCGCACT 3648
 RESULT 7
 ABX97048
 ID ABX97048 standard; cDNA; 4268 BP.
 XX
 AC ABX97048;
 XX
 DT 20-MAY-2003 (first entry)
 XX
 DE Human NOV24a cDNA.
 XX
 KM NOVX; cytostatic; cardiant; antiarteriosclerotic; antiaethmatic; cancer;
 KM hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
 KW human; gene; ss.
 OS Homo sapiens.
 XX
 PN WO200272757-A2.
 XX
 PD 19-SEP-2002.
 XX
 PF 08-MAR-2002; 2002WO-US06908.
 XX
 PR 08-MAR-2001; 2001US-274101P.
 PR 08-MAR-2001; 2001US-274194P.
 PR 08-MAR-2001; 2001US-274281P.
 PR 08-MAR-2001; 2001US-274322P.
 PR 09-MAR-2001; 2001US-274849P.
 PR 12-MAR-2001; 2001US-275235P.
 PR 13-MAR-2001; 2001US-275578P.
 PR 13-MAR-2001; 2001US-275579P.
 PR 13-MAR-2001; 2001US-275601P.
 PR 14-MAR-2001; 2001US-276000P.
 PR 16-MAR-2001; 2001US-276776P.
 PR 19-MAR-2001; 2001US-276994P.
 PR 20-MAR-2001; 2001US-277239P.
 PR 20-MAR-2001; 2001US-277321P.
 PR 20-MAR-2001; 2001US-277327P.
 PR 21-MAR-2001; 2001US-277791P.
 PR 22-MAR-2001; 2001US-277833P.
 PR 23-MAR-2001; 2001US-278152P.
 PR 26-MAR-2001; 2001US-278894P.
 PR 27-MAR-2001; 2001US-278999P.
 PR 27-MAR-2001; 2001US-279036P.
 PR 28-MAR-2001; 2001US-279344P.
 PR 30-MAR-2001; 2001US-27938P.
 PR 30-MAR-2001; 2001US-279959P.
 PR 30-MAR-2001; 2001US-280233P.
 PR 30-MAR-2001; 2001US-280822P.
 PR 02-APR-2001; 2001US-280822P.
 PR 02-APR-2001; 2001US-280822P.
 PR 02-APR-2001; 2001US-280900P.
 PR 04-APR-2001; 2001US-281194P.
 PR 13-APR-2001; 2001US-283675P.
 PR 30-APR-2001; 2001US-287424P.
 PR 02-MAY-2001; 2001US-288066P.
 PR 03-MAY-2001; 2001US-288342P.
 PR 03-MAY-2001; 2001US-288528P.
 PR 15-MAY-2001; 2001US-291190P.
 PR 16-MAY-2001; 2001US-291099P.
 PR 16-MAY-2001; 2001US-291240P.
 PR 30-MAY-2001; 2001US-294485P.
 PR 31-MAY-2001; 2001US-294889P.
 PR 31-MAY-2001; 2001US-294899P.
 PR 18-JUN-2001; 2001US-299027P.
 PR 19-JUN-2001; 2001US-299303P.
 PR 19-JUN-2001; 2001US-299310P.

QY 2085 CCAGSTCCAGCCGTTTCTGAGGACCTTACAGAGACGATGAGGGGTATGAGGACTGGCT 2144
 Db 1980 ACAGGTCCAGCGCTTCTGAGGACCTTACAGAGACGATGAGGGGTATGAGGACTGGCT 2039
 QY 2145 GCGATGAGGAGACGAGCCCTTCTGCTTGTACACCTTACCTCATGTGATGCGCTGAGACC 2204
 Db 2040 GCGATGAGGAGACGAGCCCTTCTGCTTGTACACCTTACCTCATGTGATGCGCTGAGACC 2099
 QY 2205 TGCAGATCCAAAGTTGAGTGTGCTGTGGGGGATTTCTCTGACCGCTGGCAATGGAAATGG 2264
 Db 2100 TGCAGATCCAAAGTTGAGTGTGCTGTGGGGGATTTCTCTGACCGCTGGCAATGGAAATGG 2159
 QY 2265 CTCGAAACGAGACGACCTCCCGATGTACCAACAGATGCGGAGCTGTATACAGTGC 2324
 Db 2160 CTCGAAACGAGACGACCTCCCGATGTACCAACAGATGCGGAGCTGTATACAGTGC 2219
 QY 2325 AGGAGGGGAGGGCTCCCTGAGGAAATGTGACCACTTACCTTAAATCTTCTGGGAGTT 2381
 Db 2220 AGGAGGGGAGGGCTCCCTGAGGAAATGTGACCACTTACCTTAAATCTTCTGGGAGTT 2279
 QY 2382 TAAAGAAACCTTACAGGCTTACGCGTGGGACGCGGAGTGTGCAATGACAGAAATGCAAT 2441
 Db 2280 TAAAGAAACCTTACAGGCTTACGCGTGGGACGCGGAGTGTGCAATGACAGAAATGCAAT 2339
 QY 2442 CCTCAATCAAGCTGTTCCTCGAGCAAGAGCTGAGGATCTTATAGCCAAAGTCCAACTCT 2501
 Db 2340 CCTCAATCAAGCTGTTCCTCGAGCAAGAGCTGAGGATCTTATAGCCAAAGTCCAACTCT 2339
 QY 2502 GATGCAAGATGAAAGATGATCATAGAGTAAATTTTCCATGAGAGCTGGAAAGTCAATC 2561
 Db 2400 GATGCAAGATGAAAGATGATCATAGAGTAAATTTTCCATGAGAGCTGGAAAGTCAATC 2459
 QY 2562 AGTGTATGAGGAGGACGATTTATGTGACTATGACAGATTCGAAATCTGATTTGTC 2621
 Db 2460 AGTGTATGAGGAGGACGATTTATGTGACTATGACAGATTCGAAATCTGATTTGTC 2519
 QY 2622 AGCCAACTTTGTTGACCATCTCCGCAATGCTTGTGAGAGTCTGATAGAGGTGCCAG 2681
 Db 2520 AGCCAACTTTGTTGACCATCTCCGCAATGCTTGTGAGAGTCTGATAGAGGTGCCAG 2579
 QY 2682 AGTGTGTGACCTCGTGAATCTTCTGAAACCCCACTATCATGCGGAGGTCTTCTGGG 2741
 Db 2580 AGTGTGTGACCTCGTGAATCTTCTGAAACCCCACTATCATGCGGAGGTCTTCTGGG 2639
 QY 2742 AAACCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2801
 Db 2640 AAACCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2699
 QY 2802 GCGGAGAACTCCCAAGAGCTAGCCAGGCTGAGAGGCTTACAGCCGAGCTTACCGGAGCAG 2861
 Db 2700 GCGGAGAACTCCCAAGAGCTAGCCAGGCTGAGAGGCTTACAGCCGAGCTTACCGGAGCAG 2759
 QY 2862 CATGCGGAGCTGTGAGGAGTACGAGCGCTATGACACGACGAGAGGACTTCTGTGTGCT 2921
 Db 2760 CATGCGGAGCTGTGAGGAGTACGAGCGCTATGACACGAGAGGACTTCTGTGTGCT 2819
 QY 2922 GAGAGCTTCTTCCAGAAACATCCAGCTCTCTGTCTGTGCGAGTGGCTCCAGATACGTC 2981
 Db 2820 GAGAGCTTCTTCCAGAAACATCCAGCTCTCTGTCTGTGCGAGTGGCTCCAGATACGTC 2879
 QY 2982 CTTCCTTGGCCCGAGCTGACATCCACCAATTCAGAAATTCGATCCGAGCTGGCGAGAGC 3041
 Db 2880 CTTCCTTGGCCCGAGCTGACATCCACCAATTCAGAAATTCGATCCGAGCTGGCGAGAGC 2939
 QY 3042 CCTTTGACCAATATGCTTGAACCACTTGAAGCAAAACAGAGACCTTGAACCTGAGAGC 3101
 Db 2940 CCTTTGACCAATATGCTTGAACCACTTGAAGCAAAACAGAGACCTTGAACCTGAGAGC 2999
 QY 3102 AGGAGTCCCATACCTGTCCCACTAGATGAGGCTTCTGTGAACCCCTCGGAATAG 3161
 Db 3000 AGGAGTCCCATACCTGTCCCACTAGATGAGGCTTCTGTGAACCCCTCGGAATAG 3059
 QY 3162 TAACTACGATACCCCATCAAGCCAGCATTTGAGAACTGGGGCGATGCTTCTGTGTAC 3221

Db 3060 TAACTACGATACCCCATCAAGCCAGCATTTGAGAACTGGGGCGATGACTTCTGTGTAC 3119
 QY 3222 AGAGTGAAGGCTTCCAAATGATGTTCACACTCTGTTCACACAGCTCCGACACGACAT 3281
 Db 3120 AGAGTGAAGGCTTCCAAATGATGTTCACACTCTGTTCACACAGCTCCGACACGACAT 3179
 QY 3282 CAAAGTGTGGCGCCCTGGGTGATCTGTGATCAACGACGATGGGAGCTGACCAACAA 3341
 Db 3180 CAAAGTGTGGCGCCCTGGGTGATCTGTGATCAACGACGATGGGAGCTGACCAACAA 3239
 QY 3342 CTCAGTGACCTTACCAATCTTGAAGGAGCTCTTGTGAGCATTTGAGAGGAGATGGGAA 3401
 Db 3240 CTCAGTGACCTTACCAATCTTGAAGGAGCTCTTGTGAGCATTTGAGAGGAGATGGGAA 3299
 QY 3402 CTTGAGACTCAACCAACTGCTCCCAACTTGTGAAGAAATTCACCTTCTCTGG 3461
 Db 3300 CTTGAGACTCAACCAACTGCTCCCAACTTGTGAAGAAATTCACCTTCTCTGG 3359
 QY 3462 CTTCTTACAGGACCTGGGAGGGGAGCAGCAGGACTTAAATGTGGCAGCGGAGGGGCGAG 3521
 Db 3360 CTTCTTACAGGACCTGGGAGGGGAGCAGCAGGACTTAAATGTGGCAGCGGAGGGGCGAG 3419
 QY 3522 AGCT---AGGACATGCGACGCGCCAGGCTGGGACTGTGAGACGAAATGAAAAACAGCCC 3578
 Db 3420 AGCTAGGAGGAGCATGCGACGCGCCAGGCTGGGACTGTGAGACGAAATGAAAAACAGCCC 3479
 QY 3579 CGACATCACTGAGGAGAAAGACTGGAAGCTGTGACACTTTCATTTGGGCTCAACGACTT 3638
 Db 3480 ---CATACCTTTCAGGAAAGACTGGAAGAAATTAACCTTGTATAGGCGGCAATGACT 3536
 QY 3639 GTGTCAATCTGTGAGAAATCCGAGGCCCACTTGGCCACGAAATATGTCAGACATCCA 3698
 Db 3537 GTGTCAATCTGTGAGAAATC-----TGTAGTGAATATGTCAGACATCCA 3584
 QY 3699 ACAGGCTTGTGACATCTCTGTGAGAGCTCCCAAGGCTTTCGTCAACCTGTGTGAGGT 3758
 Db 3585 ACAGGCTTGTGACATCTCTGTGAGAGCTCCCAAGGCTTTCGTCAACCTGTGTGAGGT 3644
 QY 3759 CATGAGCTGTGAGCTGTATCCAGGCGCAAGGCGGAAATGTGCCATGTGGCAGCTCA 3818
 Db 3645 CATGAGCTGTGAGCTGTATCCAGGCGCAAGGCGGAAATGTGCCATGTGGCAGCTCA 3704
 QY 3819 GAACTGCTGACTGTGCTGAGACATCGGAAGCTCCCTGTGAGAAACGAAACTGAAAGAA 3878
 Db 3705 GAACTGCTGACTGTGCTGAGACATCGGAAGCTCCCTGTGAGAAACGAAACTGAAAGAA 3764
 QY 3879 AGTGAATGGAACCTTCAGATGAGCATCTCAGATTTCTCTTACTGTGACCAATATACACA 3938
 Db 3765 AGTGAATGGAACCTTCAGATGAGCATCTCAGATTTCTCTTACTGTGACCAATATACACA 3824
 QY 3939 GCGTGAAGACTTGTGCGGTGTGTGTGAGGCTTTCTTCAAAACACATCACCACCTGAA 3998
 Db 3825 GCGTGAAGACTTGTGCGGTGTGTGTGAGGCTTTCTTCAAAACACATCACCACCTGAA 3884
 QY 3999 CGAGAGGAGGAGACATGACTTCACTTCTTCTCCGAGAGCTGTTTCACTTCTCAAGCCG 4058
 Db 3885 C---AGAGGAGGAGACATGACTTCACTTCTTCTCCGAGAGCTGTTTCACTTCTCAAGCCG 3941
 QY 4059 CGGAGATGCGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4118
 Db 3942 CGGAGATGCGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4001
 QY 4119 GACTTACTCCAACTTCAACCCACAGCGAGGCAAACTCAAGTGGCCCTCTCTGAGAG 4178
 Db 4002 GACTTACTCCAACTTCAACCCACAGCGAGGCAAACTCAAGTGGCCCTCTCTGAGAG 4061
 QY 4179 CCTTACCTTCAACCTGTGAGAAAGCGAGTGTCTCCAGAGCTGAGGAGGAGGAGGAGGAGGAG 4238
 Db 4062 TCTTACCTTCAACCTGTGAGAAAGCGAGTGTCTCCAGAGCTGAGGAGGAGGAGGAGGAGGAG 4121
 QY 4239 CGAGGTCTCTACTGTGGCTGTCCCACTGTGACAGGAGGAGTGGGCTGTGTGGGCAATCAT 4298

Db 4122 CGAGTCTCTACTGAGGCTGTCCAGTGGCAGCGGAGTGGCCCTTGNGTGGGCATCAT 4181
Qy 4299 CGGAGACGTGCTGTGAGGTGACAGAGAGTGGCCGAGGAGATCTCCATGAGGCT 4358
Db 4182 CGGAGACGTGCTGTGAGGTGACAGAGAGTGGCCGAGGAGATCTCCATGAGGCT 4241
Qy 4359 GCGCACTGTGGCCCTCTAG 4377
Db 4242 GCGCACTGTGGCCCTCTAG 4260
RESULT 8
AAT91874
ID AAT91874 standard; cDNA; 4613 BP.
AC AAT91874;
XX
XX 28-JAN-1998 (first entry)
XX
XX Rat phospholipase-B/lipase cDNA from plasmid pSVL-RIPLB.
XX
XX phospholipase B; lipase; supplement; pancreatic phospholipase; reagent;
XX screening; rat; ds.
XX
XX Rattus rattus.
XX
XX Key Location/Qualifiers
XX CDS 36..4388
XX FT /*cag= a
XX FT sig_peptide 36..125
XX FT /*cag= b
XX FT mat_peptide 126..4385
XX FT /*cag= c
XX FT /note= "putative mature protein"
XX
XX JP09248190-A.
XX
XX PD 22-SEP-1997.
XX
XX PF 15-MAR-1996; 96JP-0086022.
XX
XX PR 15-MAR-1996; 96JP-0086022.
XX
XX PA (TOJO/) TOJO H.
XX PA (TOFU) TONEN CORP.
XX
XX WPI; 1997-520744/48.
XX DR P-PSDB; AAM30751.
XX
XX PT DNA encoding protein with phospholipase B and lipase activity -
XX PT useful as supplement to pancreatic phospholipase
XX
XX PS Claim 4; Pages 7-12; 16pp; Japanese.
XX
XX CC The present cDNA encodes a protein which has both phospholipase B and
XX CC lipase activity. The claimed novel protein comprises at least amino
XX CC acids 367-712 of AAM30751. The phospholipase is useful as a supplement
XX CC to pancreatic phospholipase and as a reagent for the determination of
XX CC phospholipase B/lipase.
XX
XX SQ Sequence 4613 BP; 1218 A; 1243 C; 1142 G; 1010 T; 0 other;
Query Match 61.4%; Score 2687; DB 18; Length 4613;
Best Local Similarity 77.1%; Pred. No. 0;
Matches 3340; Conservative 0; Mismatches 970; Indels 21; Gaps 5;

Qy 148 CCAATTAATTAGAGTGAATATGCTCTTAATTCAGTTCATCTGTGAAGCCTTGAT 207
Db 189 CCAAGAGATTAAATCAGAGTGTGCTTTTAATCAGTTCATCTGTGAAGCCTTGAC 248
Qy 208 ATTAATTTGGCAGCAGCATTTGGCAATCTGAAATTTCTTCAGAACCCAGGAGCGCG-- 265
Db 249 ATTAATTTGGCAGCAGCATTTGGCAATCTGAAATTTCTTCAGAACCCAGGCGTG 308
Qy 266 ----ATCTGGAAGAGCAAGCTGAGTGAAG--GCCACAGCAGGCTGATGGAGTG 318
Db 309 GTCAACATGAGAAACCTCAAGCCTTCGAGAGCAATGAGAAATGTGCATAGAAATC 368
Qy 319 ATGACAGCTCTTTCAGACATCATCAATATTTGATCTCTGTGTCATAGCCTGTGTGC 378
Db 369 ATGACAGCCTTTTCAGATATCATCAACATTTCAACCTTCTGTGATGCCACGTGT 428
Qy 379 CACACTGGAAGAGAGTCAATACCCAC--GATGTGTGGAAGACTTGTGATTCAGGCT 435
Db 429 TCTCTGGAAAGGATCTGAGAGCCACACTAATTTGCCAAGACTTGTGATTCAGGCT 488
Qy 436 CAAGAACTGTGAGAAACATGAAAGAAAGAACTGCACTTGAATTTGACTGGAAG 495
Db 489 AAAGAGCTGTGAGGCAATGAAAGCAACCCGAACTTGATTTGAAAGACTGGAAG 548
Qy 496 CTGATCAATGTGTTCTTCAATATGCAAGCAGTGTATCTGTGCCCTGTGCTCAACAG 555
Db 549 CTGATCACTGTGCTTTCAGATTAACAGAGCAATGCACTGTGTCTGTATCAGCAG 608
Qy 556 AATGGCTTGGCGCGCGCGCGGTGATGATGATGATGATGATGATGATGATGATGATGAT 615
Db 609 AAAGGCACTTGATGAAGCACTAGGAGATGTCTGAGGGGTGCTGATTAATCTGATCCT 668
Qy 616 GAGTCTCCAGAGCATTTTGAACCTGTGTGAGTCTTGTGAGTGTGAGAGTCTCTGCT 675
Db 669 GAGTCTCCAGAGCGTTTGTGAATTTGTGTGATCTCTGAGGTTTGAACATGAGCTCAG 728
Qy 676 CAGTATCAGGCACTTGGCTCAGCCTTCAGCAGCAGAGCCTGTAAATTTGCTCAGAGAGACC 735
Db 729 CAGATTAAGAGACTGTGTTTCAAGCCTTCAGCAGAGATTTGCAAAATCTCAGAGAAATA 788
Qy 736 ACCCGCTGGCCAAAGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 795
Db 789 ACGAAGTTATCCAAAGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 848
Qy 796 GCTTCCAGCAGGTACAGTGAAGAGAGTCTTCAACGTTGTTTCCAGCCTTCTTAT 855
Db 849 GCCTCCAGCAAGTTCAATGAAGATGAGACCTTGTGTGTGTTTCCAGTCTTCTCTCT 908
Qy 856 GAGACCAACCCCATCTCTACCTGAGAGAGACCCCGACTCCAGGATTTACACGCTGGCC 915
Db 909 GAG-----GTAAATCTACCTTGTGAAGAGCCTTCGCCCAAGATTCACACGCTGCC 962
Qy 916 TGGCATCTCTGAATAGATGATGAGCCAGCAGAGAGAAAGATGAGCCATTTGATGATGTA 975
Db 963 CTGAGATTTGAAATGATGATGAGAAACAGTGGTGAAGAGATGGAACATCAATGA 1022
Qy 976 AAACACGGAGGCGCAATGAAGTCTCTTCAGAGAGAGCCCTTATCTGTTCAGTACAGA 1035
Db 1023 GCAAGAGAGAAACCAATGAATGTCTCTCAGAGAGAGCCCTTATCTGTTCAGTACAGA 1082
Qy 1036 AAGACCAACTACCTGAGCCAGACCTGAGAAACCCCAAGACCAAGCTTGAAGTGAAGAGA 1095
Db 1083 AATAGCAACTACCAAGCCAGACCTGAGAAACCCCAATGAGAAATTTTCAATGAAGAGAGA 1142
Qy 1096 GCGGAATCAGATGTCTCTGCAAAAGACCCCTTCGATACGTTCCACTGATTCATAGG 1155
Db 1143 ACAAATTTACCTGTCTGCAAAAGACCCCTGGAATCCATCCCAACAGTTCAAGG 1202
Qy 1156 CTGAAGCGGCTGACATCAAGTGAATTTGAGCCCTGGGTGATCTCTCAGCGCAGGCAAT 1215
Db 1203 CTGAGCGCGCTGACATCAAGTGAATTTGAGCCCTGGGTGATCTCTCAGCGCAGGCAAT 1262

QY 1216 GGGGCGGGGTGCAACCTTGGAAAGTCTTGGACGCTTGACTGACGAGCCGCTGTC 1275
DB 1263 GGGGACAGGGGTCCAGCCCTGGGAATGCTTGGATGCTTAACTCAATPCCAGAGCCGTGTG 1322
QY 1276 TGGAGGTGGGCGGGAATGAAACATGCGACCCGTTACCAACCTGGCGAAACATCTCCGG 1335
DB 1323 TGGAGTGGGCGGGAATGAAACATGAGACCGTGAACCCGTGACCAACCTGAGCAACATCTCCGG 1382
QY 1336 GAATTAACCCCTTCCCTGAAAGGGCTTCTGTTGGCACTGGGAAGAAACAGTCCTAAT 1395
DB 1383 GAATTAACCCCTTCTGTAAGGGCTTCTGTTGGCACTGGGAAGAAACATCTCCCGA 1442
QY 1396 GCGCTTTAAACCAAGCGCTGTGCGAGAGCCGAGCTGAGGATCTACCTGTCCAGGCGAG 1455
DB 1443 GCATCTTTCAACCAAGCGCGTGAAGAGGCAAAATCTGATGGCTTAAGTGGCCAGGCCAA 1502
QY 1456 AGGCTGTGACCTGATGAAGAAATGACACAGAGATACCTTTAGGAAGAGCTGGAAGATA 1515
DB 1503 AAGCTGTGAGCTGTGATGAAGGATGACAGAGCAATAACTTTCAGGAGAGCTGGAAGATA 1562
QY 1516 ATAAACCTGTTTATAGGCGGCAATGACCTGTGATTTTGCATATGATCTGTCCACTAT 1575
DB 1563 ATACATGTTTATAGAGGCAACGACCTGTGAGCTTCTGCAATTAACCTGCGCTGCTT 1622
QY 1576 TCTCCCGAAGCTTCAAGACAACTTGAAGAGCCCTGAGCAATCTCATGCTGAGGT 1635
DB 1623 TCTCCCGAAGCTTCAAGACAACTTGAAGAGCCCTGAGCAATCTCATGAGAGTT 1682
QY 1636 CTTGCGGCAATTTGTGAACCTGTGAGCGGTCTTGAATGCTCAACCTGAGGAGCTGTAC 1695
DB 1683 CCGCGGCGCTTGTGAACATGTGTGCTGATGAGATCAACCCCTTGAAGAACTGTTC 1742
QY 1696 CAGGAGAAAAAGTCACTGCGCCCAAGAGATGATCTCAAGTCTGTGCTCTGTGCTTC 1755
DB 1743 AATGAACCTTAAGTCAAGCTGCGCCCAAGAGATCTCAAGAGCTGTGCTGTGCTTC 1802
QY 1756 AAGTTGATGATTAACCAAGAACTTGTCTAACCTCATCGAATTCACAAAGAGTTTCAG 1815
DB 1803 AACCTTGTGAGAACTCAAGAGAACTTGTCCCACTGTGGAAGAGAAAGAGAGATTCAG 1862
QY 1816 GAGAGAACCAACCACTGATTTGAGAGTGGCGATATGACAAAGGAGATTTACTGTG 1875
DB 1863 GAAAGAACTGGAAGAACTGATTTGAGAGTGGCGATATGACAAAGAGATTCACCGTG 1922
QY 1876 GTTGTGACGCGCTTCTTGAAGACGTGAGACATGCGAAAGACCTCGAGAGATTCCTGAC 1935
DB 1923 GTCTCTCAGCCCATGTTTGAAGATGTGTCATGCCAGGACCTGTGAGGCTTGTCCGAC 1982
QY 1936 AACTCTTCTTCTGCTCTGACTGTGTTTCACTTCAAGAGCAAGTCTCACTCCGAGAGCC 1995
DB 1983 AGCTCTTCTTCTGCTCTGACTGTGTTTCACTTCAAGTCAAGACTCAAGCTCCGCTAGCC 2042
QY 1996 AGTGTCTCTGGAACAATATCTGAGACCTGTTGGCCAGAGAGCACTGCTCAATAGTTT 2055
DB 2043 ATGCGCTCTGGAAGAAACATGTAAGAACTGTGGGCGGAAAGCAAGACATCAAGATTTT 2102
QY 2056 GAAAAAAGATCAATATCAATGTCGAGACCAAGTCCAGCGCTTCTGAGAGACTTCAAG 2115
DB 2103 GAAATCAAGGTCCCTATCATGTGTCCAAACCAAGACTCAAGCTTCTGAGCAACCAAG 2162
QY 2116 AACAGCATGAGAGGATATGGGACCTGGCGTCCATGACGAGAGAGAGCCCTTGTGCTTG 2175
DB 2163 AACAGCACTTGGAGACATGGAATCTTGATGCTTGTGAGAGAGAGAGCCCTTGTGCTCA 2222
QY 2176 CAACCTTACCTCACTGATGACCTGAGACCTGAGACATCCAAAGTGTGCTGTGAGG 2235
DB 2223 CCACCAACCTCACTGATGATACCTGAGACCTGAGACATCCAAAGTGTGAGGCTGTGAG 2282
QY 2236 GATTTCTGACCGCTGCAATGGAATGGTCCAAACAGAGACACTCCCGATGTAC 2295
DB 2283 GACTCTGTGATCTGTGCAATGGAATCAGCTCCCAAGAGATGATCTGTGATGTAC 2342
QY 2296 ACACAGTATCGGGAATGTGATACATGACAGAGGGGAGCGCTCCGTGAGATGTGAC 2355

DB 2343 ACACAGTATCGAGACCTGTCTCAAGTGTGAGGGAGAGAGTCTCTGAGAGATGTGAC 2402
QY 2356 ACCTTACCTAATATCTTGGAGGATTAACAGAAACCTCAAGGCTTAACCGCGGCGACG 2415
DB 2403 ACCTTGGCCAACTCTCCGGGAATTTAATGAAATCTCAAGCTTCACTGATGGAACC 2462
QY 2416 GGTGATGCCAATGACAGAGATGATCTCAATCAAGCTGTTCGCGAGCAAGGCTGAG 2475
DB 2463 GGTGATGCCAATCTGCAAGCGGCTTCTTAAACAGGCTGTCTCTGGCGCAAGGCTGAG 2522
QY 2476 GATCTTATGAGCCCAAGTCCAACTGTATGACAGAGATGAAGAGATGATCATAGATTAAT 2535
DB 2523 AACCTTGCAAGTCAAGTCCAGACTGTGATTCAGAAATGAAGATGACACAGAGTGAAC 2582
QY 2536 TTTCATGAAGAGCTGGAAGGTCACTACAGTGTGATGAGAGGACGCGATTTATGTACTAC 2595
DB 2583 TTTCACCAAGAGCTGGAAGGTCACTACTGTGATGATGAGAGGCGACGACTGTGTGACTTC 2642
QY 2596 TGACAGATTTGGAATGTATCTGTGAGGCAACTTGTGACCACTCCGCAATGCGCTTG 2655
DB 2643 TGCAGAGATTTGGAATGTATCTGTGAGGCAACTTGTGACCACTCCGCAATGCGCTTG 2702
QY 2656 GAGTCTTCATAGAGAGGTGCGCAGAGTCTGTGCAACCTCGTGAATCTTCTGAACCC 2715
DB 2703 GACATCTCATAGAGAGGTACCCAGAGCCGTGTCAACCTGTGACATTCATGAACCC 2762
QY 2716 ACTATATCGCGGAGGTCTTCTGAGAAACCCAGACAGAGTCCCATGACGAGCCAGC 2775
DB 2763 AGTATATTCGCGAAGTGTCTTGAAGAAACCCAGACAGAGTCCCTGTGATGACAGCAGT 2822
QY 2776 GTTGTGTATGATCGGCTTGTGACCCGCGGAGAACTCCCAAGAGTACGAGGCTGAG 2835
DB 2823 GTCTGTGCACTGTGTGTGACCCAGGAGAGATTCATATAGCTGTGAGAGGCTTGTGAG 2882
QY 2836 GCGTTCAGCGGAGCTTACCGGAGCAGCATGCGGAGCTGTGTGAGGCTCAGGCGCTATGAC 2895
DB 2883 GCGTTCAGCAATTCATCAGAGATGACATGCTTCAACTGTGTGAGTCAAGGCGCTATGAC 2942
QY 2896 AGCAGAGAGACTTCTGTGTGTGCTGACGCGCTTCTTCAAGACATTCAGACTCCGTTC 2955
DB 2943 ACCCGGAGAGATTTCTGTGTGATGACGAGCCCTTCTTCAACATTCAGAGCTCCCATC 3002
QY 2956 CTGGCGAGAGGCTCCCAATACGTCTCTTGTGCGGAGAGCTGATCCCAATTCAG 3015
DB 3003 CTAGAGAAATGGAATTCATATCATCTTCTTGTGCGGAGAGCTCTCTTACACCAAG 3062
QY 3016 AATTCACCTCCAGCTGAGCAGAGCCCTTGTGACCAATATGCTTGAACCACTTGTGAAC 3075
DB 3063 AAGTTCACATCAGCTGCGAGAGAGCCCTTGTGCGCAATATGCTTGAACCCCTGTGAAAG 3122
QY 3076 AAAACAGAGACCTGTGACCTGAGAGCAGATGCCATGACCTGTGCCATCGAATGAG 3135
DB 3123 AAAATGATATCTTGGACCCGAAAGAACTATATGCTTGTGCGCTGCCCAACAGAGCAAG 3182
QY 3136 CCGTTCGAGAAACCCCTGGAATATGATACATGATACCCCATGAGCCAGCATTTGAG 3195
DB 3183 CCGTTCGAGAAACCTTGTGGAACAGATATCACTACCTATTAAGCAGGCAATTTGAG 3242
QY 3196 AACTGGGAGAGTACTTCTGTGTACAGAGTGAAGGCTTCCAAATGTGTTCCAACTCT 3255
DB 3243 AATTTGGGAGAGTACTTCTGTGTACAGAGAGAGTCTTCCAGCAAGATACCACTCA 3302
QY 3256 GTTCACAGCTTCGACAGCAGACATCAAGATGTGTGAGCGCCCTGTGAGTACTTGTGACT 3315
DB 3303 GTTCATGAGCTTCGACATCAAGACATCAAGATGTGTGAGCGCAATGTGTGACTTGTGAG 3362
QY 3316 ACAGAGTGGAGCTTGAACCAACATCTCAATGATCTTACCAATCTTGTGAGGAGGAGCTC 3375
DB 3363 ACAGGACCGGAGCTTGAACCAAGTGTGACACAGTGTGAGACACCCCTGTGAGGAGGAGCTC 3422
QY 3376 TCTTGGAGCATTTGAGGAGGATGGGAACTTGAAGACTCAACCAACTGTGCCCAATTTCTG 3435


```

Db 3423 TCTTGAGCATTTGAGGAGATGGAACGTTGAGAACCCATACCACTTGCCCAATCTCTG 3482
Qy 3436 AAGAACTTCAACCTTACCTCTTGTGCTTCTTACGACACTGGAGGGGAGACAGAGA 3495
Db 3483 AAGAACTTCAACCTTCTTCACTCTTGTGATTTCTTCAACCGGATCCCTGGGAAACAGGAGGA 3542
Qy 3496 CTAAATGTGACAGGAGGGGCGAGAGCTAGGAGACATGCGACCCAGGCTGGGACCTG 3555
Db 3543 TTAAATGTGACAGGAGGGGCGAGAGCTAGGAGACATGCGGCGCCAGGCTCAGGCTCTG 3602
Qy 3556 GTAGAGGAGATGAAAAACAGCCCGACATCACTTGGAGAAAGACTGGAGCTGTGTACA 3615
Db 3603 GTAGAGGAGATGAAAAACAGCCCTTACCAATCAACATACAGAGAGCTGAAAGCTGATTA 3662
Qy 3616 CTCTCTCATTTGGGGGTCAACGACTTGTCTTATCTGTAGAGATCCGGAGGCGCACTTGG 3675
Db 3663 CTCTCTCATTTGGGAAACAAGAGCTGTGTCTTACTGTAGAGATCCGGAGAACTACTCAAC 3722
Qy 3676 ACGGAATATGTTGACGACATTCACAGAGCCCTGGACATCTCTCTGAGAGCTCCCAAG 3735
Db 3723 AGGAGATATGTCAAGTATCATTCAGCATGCTTGGACATCTTATAGAGAGCTTCCAGG 3782
Qy 3736 GCTTTCTCAACGCTGTGAGAGCTGAGAGCTGAGCTGTAGCTGTACAGGCGCAAGCGGG 3795
Db 3783 GTTTTCATCAACGCTGTGAGAGCTGAGAGCTGTGCTGTCTCCAGACAGAGGCGGG 3842
Qy 3796 AAATGTGCGCAGC---TGAGAGCTCAGAACATGCACTTGGCTCAGACACTGCGCAAGC 3852
Db 3843 AAATGTGCGCAGCCTTGTGGCTGTCTCAGAAAAACGTCAGCTTAAACGCTTCAAAAC 3902
Qy 3853 TCCCTGAGAGAGCAAGAACTGAGAAAGTGAAGTGAAGCTCCAGATGAGCATTCAGT 3912
Db 3903 CTATGCGCATGACAGAGAGCTGAGAAAGTCAACGGAACCTCCAGATGCTCTCGAG 3962
Qy 3913 TTCTCTACTGCGACCAATACACAGCGGTGAGAGCTTTGGCGTTGTGTGACAGCTTTC 3972
Db 3963 CTCTCTACTGCGACCTGCTCATGCAAGCTGAGAGACTTTGCAAGTCACTGTCCAGCCTTTC 4022
Qy 3973 TTCCAAACACACTCACCCGACGAGAGAGAGGGGACACTGACCTCACCTTCTTCTCC 4032
Db 4023 TTCCGAAATCCTTGTCCACTGAGATGAGGAGGGGCTCGACCTCACCTTCTTCTCT 4082
Qy 4033 GAGGACTGTTTTCATCTCTGAGACCGGCGGATCCGAGATGCGCATTCGACACTTGGAG 4092
Db 4083 GAGGACTGTTTTCATCTCTGAGACCGGCGGATCCGAGATGCGCATTCGACACTTGGAG 4142
Qy 4093 AACATGCTGAAACAGTGGGCGGCAAGACTTACCTCCAACTTCCACCCAGCGGAGCC 4152
Db 4143 AACATGCTGAAACAGTGGGCGGCAAGAACTTCAATTAATTCACATTAACCCGAGCC 4202
Qy 4153 AAATCTCAAGTGGCCCTCTCTGAGAGCCCTTAACTTCAACCCCTGGGAAACAGCCGAT 4212
Db 4203 AAATCTCAAGTGGCCCTCTCTGAGAGCCCTTAACTTCAACCTGTGCGAAACAGCAAT 4262
Qy 4213 CTCCCAACAGGCTGTGAGAGAGCCCGGAGGCTCTTACTGAGGCTGTGCCAGTGGACGG 4272
Db 4263 CTCTTAACAGAGCTTAAGAAAGAACTCAATTAACCTTCACTGAGGAGTGCAGTGGCTG 4322
Qy 4273 GGAAGTGGGCTTGTGTGAGGATCATCGGAGCACTGTGTGAGAGTGCAGAGAGAGTGGC 4332
Db 4323 GTAGAGTGGCTGTGTGAGGATCATCGGAGCACTGTGTGAGAGAGTGTGTGAGAGTGTG 4382
Qy 4333 CGAGGAGAGAG 4343
Db 4383 CAATTAAGAG 4393

```

```

DT 16-DEC-2002 (first entry)
XX
XX Human phospholipase protein encoding cDNA SEQ ID NO:1.
DE
XX Human, phospholipase, enzyme; chromosome 2; gene; ss.
KM
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT 5'UTR 1..175
FT CDS /*tag= a
FT /*tag= b
FT /*tag= c
FT 3'UTR /product= "phospholipase"
FT /*tag= c
PN WO200262977-A2.
XX
XX 15-AUG-2002.
PD
XX
XX 28-JAN-2002; 2002MO-US02302.
PF
XX 08-FEB-2001; 2001US-0778961.
PR
XX (PEKE ) PE CORP NY.
PA
XX
XX Van C, Ketchum KA, Di Francesco V, Beasley EM,
PI
XX WPI; 2002-682698/73.
DR
XX P-PSDB; ABP53556.
XX
XX New human phospholipase proteins, useful for the development of human
PT therapeutics and diagnostic compositions, drug screening assays, tissue
PT typing and pharmacogenomic analysis
PS
XX Claim 4; Fig 1A; 95pp; English.
XX
XX The present sequence encodes a human phospholipase protein (1) located on
CC chromosome 2. (1) can be used for identifying agents that modulate its
CC function or activity where the agent is useful for treating a disease
CC or condition mediated by a the human phospholipase protein. (1) peptides
CC can be used in substantial and specific assays related to functional
CC information of the peptide sequences, to raise antibodies or to elicit
CC immune response, as reagents in assays that determine the levels of
CC protein in biological fluids, and as markers for tissues where the
CC corresponding protein is expressed. Nucleotide sequences encoding (1)
CC can be used as probes, primers and chemical intermediates in biological
CC assays, for constructing recombinant vectors, and expressing antigenic
CC portions of the protein. (1) and nucleic acid molecules encoding it can
CC be used in the identification of therapeutic proteins and may serve as
CC models or targets for the development of human therapeutic agents that
CC modulate phospholipase activity in cells and tissues that express the
CC phospholipase, such as in kidney, blood, lung, brain glioblastomas,
CC prostate, colon or leukocytes.
XX
XX Sequence 1835 BP; 458 A; 549 C; 474 G; 354 T; 0 other;
SO

```

```

Query Match 31.1%; Score 1362; DB 24; Length 1835;
Best Local Similarity 89.3%; Pred. No. 0;
Matches 1582; Conservative 0; Mismatches 0; Indels 190; Gaps 3;

```

```

Qy 2618 CTGAGCGCACTTTGTGACCATCTCCGCAATGCTTGGACGTCCTGCAATAGAG----- 2671
Db 1 CTGAGCGCACTTTGTGACCATCTCCGCAATGCTTGGACGTCCTGCAATAGAGACTTT 60
Qy 2672 -----AGTGGCCAGAGTCTGTCACCTCGTGAACCTTCTGAAACCCCACTATATGC 2725
Db 61 TCCCTTAGTGGCCAGAGTCTGTCACCTCGTGAACCTTCTGAAACCCCACTATATATGC 120
Qy 2726 GGCAGGTGTTCTGTGGAAACCCAGACAGTGCCTGAGCAGAGCCAGCGCTTTGTGTA 2785
Db 121 GGCAGGTGTTCTGTGGAAACCCAGACAGTGCCTGAGCAGAGCGCA----- 168

```


QY 2786 ACTGCGTTGACCTGCGGGAGAACTCCCAAGAGTAGCCAGGCTGGAGGCTTCAGCC 2845
 Db 169 ----- 168
 QY 2846 GAGCTTACCGGAGCAGATCCGCGAGCTGTGGGGTCAGGCCCGCTATGACAGCAGAGG 2905
 Db 169 -----GAGCAGCATCGCGAGCTGTGGGGTCAGGCCCGCTATGACAGCAGAGG 218
 QY 2906 ACTTCTGTGAGTGTGACAGCCCTTCTTCAGAACTCCAGAGCTCCCTGTCTGGCGATG 2965
 Db 219 ACTTCTGTGAGTGTGACAGCCCTTCTTCAGAACTCCAGAGCTCCCTGTCTGGCGATG 273
 QY 2966 GGTCTCCAGATAGCTCTCTTTGGCCCAAGTGCATCCACCAATCAGAAATTCACCT 3025
 Db 274 ----- 273
 QY 3026 CCCAGCTGGCAGAGCCCTTTGAGCAAAATGCTTGAACCACTTGGAAAGCAAAACAGAGA 3085
 Db 274 -----GCTTGAACCACTTGGAAAGCAAAACAGAGA 302
 QY 3086 CCCTGAGACCTGAGAGAGAGATGCCATTCACCTGTCCCACTCAGAAATGAGCCCTTCTGA 3145
 Db 303 CCCTGAGACCTGAGAGAGAGATGCCATTCACCTGTCCCACTCAGAAATGAGCCCTTCTGA 362
 QY 3146 GAACCCCTCGGAATGATATACACGATACCCCATCAAGCCAGCCATTGAGAACTGGGGCA 3205
 Db 363 GAACCCCTCGGAATGATATACACGATACCCCATCAAGCCAGCCATTGAGAACTGGGGCA 422
 QY 3206 GTGACTTCTGTGTACAGAGTGAAGGCTTCAATAGTGTTCACCTGTCTGACAGC 3265
 Db 423 GTGACTTCTGTGTACAGAGTGAAGGCTTCAATAGTGTTCACCTGTCTGACAGC 482
 QY 3266 TCCGACCAAGAGCATCAAAAGTGTGCGCCCTGGGTGACTCTTGACTACAGCAGTGG 3325
 Db 483 TCCGACCAAGAGCATCAAAAGTGTGCGCCCTGGGTGACTCTTGACTACAGCAGTGG 542
 QY 3326 GAGCTGACCAAAACAATCCAGTACCTACCAATCTTGGAGGGGACTCTTGGAGCA 3385
 Db 543 GAGCTGACCAAAACAATCCAGTACCTACCAATCTTGGAGGGGACTCTTGGAGCA 602
 QY 3386 TTGGAGGGGATGGGAACTTGGAGACTCACACACACTGCCCAATCTTGAAGAATGCA 3445
 Db 603 TTGGAGGGGATGGGAACTTGGAGACTCACACACACTGCCCAATCTTGAAGAATGCA 662
 QY 3446 ACCCTTACTCTTGGCTTCTTACCAAGCACTTGGAGGGGAGACAGAGACTAATGTGG 3505
 Db 663 ACCCTTACTCTTGGCTTCTTACCAAGCACTTGGAGGGGAGACAGAGACTAATGTGG 722
 QY 3506 CAGCGGAGGGGACAGAGCTAGGAGACATGCCAGGCCCTGGGACCTGTGAGACGAA 3565
 Db 723 CAGCGGAGGGGACAGAGCTAGGAGACATGCCAGGCCCTGGGACCTGTGAGACGAA 782
 QY 3566 TGAATAAAGAGCCCGACATCAACCTGGAGAAAGACTGGAAGTGTCAACTCTTCAATG 3625
 Db 783 TGAATAAAGAGCCCGACATCAACCTGGAGAAAGACTGGAAGTGTCAACTCTTCAATG 842
 QY 3626 GGGTCAACGACTTGTGTCACTTATCTGTAGAAATCCGAGGGCCCACTTGGCCAGGAATATG 3685
 Db 843 GGGTCAACGACTTGTGTCACTTATCTGTAGAAATCCGAGGGCCCACTTGGCCAGGAATATG 902
 QY 3686 TTCAGACATCCCAACAGGCCCTGAGACATCTCTTGAAGAGCTCCCAAGGGCTTTGTCTCA 3745
 Db 903 TTCAGACATCCCAACAGGCCCTGAGACATCTCTTGAAGAGCTCCCAAGGGCTTTGTCTCA 962
 QY 3746 ACCTGTGAGAGTCAATGAGAGCTGAGCTGATACCAAGGGCCAAAGGGGAAATGTGCCA 3805
 Db 963 ACCTGTGAGAGTCAATGAGAGCTGAGCTGATACCAAGGGCCAAAGGGGAAATGTGCCA 1022
 QY 3806 TGTGTGAGAGTCAAACTGCACTTGTGCTTCAAGCACTTGCAGAAAGCTTCTGGAGAAAGC 3865
 Db 1023 TGTGTGAGAGTCAAACTGCACTTGTGCTTCAAGCACTTGCAGAAAGCTTCTGGAGAAAGC 1082

QY 3866 AAGACTGAAGAAAGTGAACCTGGAACCTCAGCATCTCCAGTCTTCTTCTACTGCG 3925
 Db 1083 AAGACTGAAGAAAGTGAACCTGGAACCTCAGCATCTCCAGTCTTCTTCTACTGCG 1142
 QY 3926 ACCAATPACACAGCGGTGAGACTTTGGGTGTGTGACAGCTTTCTTCAAAACACAC 3985
 Db 1143 ACCAATPACACAGCGGTGAGACTTTGGGTGTGTGACAGCTTTCTTCAAAACACAC 1202
 QY 3986 TCACCCCACTGAACGAGAGAGGGGACATGACCTCACTTCTTCTCGAGGACTGTTTC 4045
 Db 1203 TCACCCCACTGAACGAGAGAGGGGACATGACCTCACTTCTTCTCGAGGACTGTTTC 1262
 QY 4046 ACTTCTCAGACCGCGGGACATGCCAGATGCGCATCGCACTTGAACCAATGCTGAAAC 4105
 Db 1263 ACTTCTCAGACCGCGGGACATGCCAGATGCGCATCGCACTTGAACCAATGCTGAAAC 1322
 QY 4106 CAGTGGCGCGAAGACTACCTCCACAACTTCAACCCACAGCCGACCAACTCAAGTGC 4165
 Db 1323 CAGTGGCGCGAAGACTACCTCCACAACTTCAACCCACAGCCGACCAACTCAAGTGC 1382
 QY 4166 CCTCTCTGAGAGCCCTTACCTTACACCTGCGGAAACGCGGATGCTCCAGACAGG 4225
 Db 1383 CCTCTCTGAGAGCCCTTACCTTACACCTGCGGAAACGCGGATGCTCCAGACAGG 1442
 QY 4226 CTGAAGAAGCCCGAGAGTGTCTTACTGTGGGCTGTCCAGTGGCAGCGGAGTGGCTTG 4285
 Db 1443 CTGAAGAAGCCCGAGAGTGTCTTACTGTGGGCTGTCCAGTGGCAGCGGAGTGGCTTG 1502
 QY 4286 TGTGTGGCATCATCGGAGCAGTGTGTGAGAGTGCAGAGAGGTGGCCGAGGGAAGATC 4345
 Db 1503 TGTGTGGCATCATCGGAGCAGTGTGTGAGAGTGCAGAGAGGTGGCCGAGGGAAGATC 1562
 QY 4346 CTCCAATGAGCTGCGGCATCTGCGCCCTTAG 4377
 Db 1563 CTCCAATGAGCTGCGGCATCTGCGCCCTTAG 1594
 RESULT 10
 ABX97049
 ID ABX97049 standard; cDNA, 1624 BP.
 XX
 AC ABX97049;
 XX
 DT 20-MAY-2003 (first entry)
 XX
 XX Human NOV24b cDNA.
 DE
 XX NOVX; cytostatic; cardiatic; antiarteriosclerotic; antiasthmatic; cancer;
 KM hypotensive; cardiomyopathy; bronchial asthma; gene therapy; vaccine;
 KM human; gene; ss.
 OS Homo sapiens.
 XX
 PN MO200272757-A2.
 XX
 PD 19-SEP-2002.
 XX
 XX 08-MAR-2002; 2002MO-US06908.
 PF
 XX 08-MAR-2001; 2001US-274101P.
 PR 08-MAR-2001; 2001US-274194P.
 PR 08-MAR-2001; 2001US-274281P.
 PR 08-MAR-2001; 2001US-274322P.
 PR 09-MAR-2001; 2001US-274849P.
 PR 12-MAR-2001; 2001US-275235P.
 PR 13-MAR-2001; 2001US-275578P.
 PR 13-MAR-2001; 2001US-275579P.
 PR 13-MAR-2001; 2001US-275601P.
 PR 14-MAR-2001; 2001US-276000P.
 PR 16-MAR-2001; 2001US-276776P.
 PR 19-MAR-2001; 2001US-276994P.
 PR 20-MAR-2001; 2001US-277239P.
 PR 20-MAR-2001; 2001US-277321P.

PR 20-MAR-2001; 2001US-277327P.
 PR 21-MAR-2001; 2001US-277791P.
 PR 22-MAR-2001; 2001US-277833P.
 PR 23-MAR-2001; 2001US-278152P.
 PR 25-MAR-2001; 2001US-278899P.
 PR 27-MAR-2001; 2001US-278999P.
 PR 28-MAR-2001; 2001US-279344P.
 PR 30-MAR-2001; 2001US-279338P.
 PR 30-MAR-2001; 2001US-279995P.
 PR 30-MAR-2001; 2001US-280233P.
 PR 02-APR-2001; 2001US-280802P.
 PR 02-APR-2001; 2001US-280822P.
 PR 04-APR-2001; 2001US-280900P.
 PR 13-APR-2001; 2001US-281194P.
 PR 30-APR-2001; 2001US-283675P.
 PR 02-MAY-2001; 2001US-284242P.
 PR 03-MAY-2001; 2001US-288066P.
 PR 03-MAY-2001; 2001US-288342P.
 PR 15-MAY-2001; 2001US-288528P.
 PR 16-MAY-2001; 2001US-291190P.
 PR 16-MAY-2001; 2001US-291099P.
 PR 16-MAY-2001; 2001US-291240P.
 PR 30-MAY-2001; 2001US-294485P.
 PR 31-MAY-2001; 2001US-294889P.
 PR 31-MAY-2001; 2001US-294899P.
 PR 18-JUN-2001; 2001US-299027P.
 PR 19-JUN-2001; 2001US-299303P.
 PR 19-JUN-2001; 2001US-299310P.
 PR 10-JUL-2001; 2001US-304354P.
 PR 31-JUL-2001; 2001US-309198P.
 PR 16-AUG-2001; 2001US-312903P.
 PR 10-SEP-2001; 2001US-318462P.
 PR 12-SEP-2001; 2001US-318770P.
 PR 27-SEP-2001; 2001US-325430P.
 PR 27-SEP-2001; 2001US-325681P.
 PR 18-OCT-2001; 2001US-330380P.
 PR 31-OCT-2001; 2001US-335301P.
 PR 14-NOV-2001; 2001US-332172P.
 PR 14-NOV-2001; 2001US-332271P.
 PR 14-NOV-2001; 2001US-332272P.
 PR 14-NOV-2001; 2001US-332184P.
 PR 14-NOV-2001; 2001US-332272P.
 PR 21-NOV-2001; 2001US-332094P.
 PR 03-DEC-2001; 2001US-337426P.
 PR 03-DEC-2001; 2001US-338092P.
 PR 04-DEC-2001; 2001US-337185P.
 PR 03-JAN-2002; 2002US-345705P.
 PR 07-MAR-2002; 2002US-0092900.

(CURA-) CURAGEN CORP.

PA Pediaru M, Spytek KA, Shenoy SG, Taupier RJ, Pena CE, Li L;
 XX Zernhusen BD, Gusev V, Ji W, Gorman L, Miller CE, Kekuda R;
 PI Palturajan M, Gangolli E, Vernet CAM, Guo X, Tchervnev V;
 PI Fernandes ER, Casman SJ, Malyankar UM, Gerlach V, Liu Y;
 PI Anderson D, Spederna SK, Catterton E, Burgess C, Lette M, Zhong H;
 PI Alsebrook JP, Lepley DM, Rieger DK;
 XX WPI; 2002-723332/78.
 DR P-PSDB; ABU65082.

PT NOXV polypeptides and polynucleotides, useful for preventing or
 PT treating a disorder associated with aberrant NOXV expression or
 PT activity e.g., cancer, hypertension, atherosclerosis, cardiomyopathy or
 PT bronchial asthma

PS Claim 13; Page 169; 1103pp; English.

CC This invention describes novel human NOXV polypeptides which have
 CC cytostatic, cardiant, antiarteriosclerotic, antiasthmatic and
 CC hypotensive activity. Pharmaceutical compositions comprising the NOXV
 CC proteins or nucleic acid molecules or NOXV antibodies are useful for

CC preventing or treating a disorder associated with aberrant NOXV
 CC expression or activity e.g. cancer, hypertension, atherosclerosis,
 CC cardiomyopathy or bronchial asthma. The products of the invention can
 CC be used for gene therapy or in a vaccine. ABX97008-ABX97185 are cDNA
 CC fragments amplified and isolated by the PCR primers and probes
 CC represented in ABX13460-ABX13462 and ABX97186-ABX97593. ABX97008-ABX97185
 CC encode the NOXV proteins described in ABU65041-ABU65218.

XX Sequence 1624 BP; 422 A; 421 C; 401 G; 380 T; 0 other;

SO Query Match 26.5%; Score 1160.6; DB 24; Length 1624;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1163; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy	1161	GGCGGCTGACATGAATGTAATGAGCCCTGGTGA	CTCTCTGACGCGAGGCAATGGGAGC	1220
Db	1	GCCGGCTGACATGAATGTAATGAGCCCTGGTGA	CTCTCTGACGCGAGGCAATGGGAGC	60
Qy	1221	CGGGTCCACACCTGGGACGCTTGGACGCTTGA	CTCACTACCGAGGCTCTCTGGAG	1280
Db	61	CGGGTCCACACCTGGGACGCTTGGACGCTTGA	CTCACTACCGAGGCTCTCTGGAG	120
Qy	1281	CGTCGGGAGATGAGACATCGGCACTGTACCA	CCCTGGCCAACTCTCCGGAAAT	1340
Db	121	CGTCGGGAGATGAGACATCGGCACTGTACCA	CCCTGGCCAACTCTCCGGAAAT	180
Qy	1341	CAACCTTCCCTGAAGGGCTTCTCTGTGGCA	CTGGGAAAGAAACAGCTCTAATGCTT	1400
Db	181	CAACCTTCCCTGAAGGGCTTCTCTGTGGCA	CTGGGAAAGAAACAGCTCTAATGCTT	240
Qy	1401	CTTAAACAGGCTGTGGCAGAGAGCCGAGCTGA	GTACTAATCTGTCCAGGCCAGAGGCT	1460
Db	241	CTTAAACAGGCTGTGGCAGAGAGCCGAGCTGA	GTACTAATCTGTCCAGGCCAGAGGCT	300
Qy	1461	GGTGAACCTGATGAAGATGACAGAGATACCT	TTCAGAGAACTCGAAGTAATTAAC	1520
Db	301	GGTGAACCTGATGAAGATGACAGAGATACCT	TTCAGAGAACTCGAAGTAATTAAC	360
Qy	1521	CGTGTATATAGGGGCAATGAGCTCTGTGAT	TTCAGATGATCTGTGATCTCTCC	1580
Db	361	CGTGTATATAGGGGCAATGAGCTCTGTGAT	TTCAGATGATCTGTGATCTCTCC	420
Qy	1581	CCAGAACTTCAACAGAACATTTGAAAGGCCCT	GTGACATCTTCATGAGTTCCTCG	1640
Db	421	CCAGAACTTCAACAGAACATTTGAAAGGCCCT	GTGACATCTTCATGAGTTCCTCG	480
Qy	1641	GGCATTGTGAACCTGCTGACGCTGTGATGAT	CGTCAACTTGAAGGAGCTGTACAGGA	1700
Db	481	GGCATTGTGAACCTGCTGACGCTGTGATGAT	CGTCAACTTGAAGGAGCTGTACAGGA	540
Qy	1701	GAAAAAGTCTAGTCCCAAGATGATGCTCA	GAGTCTGTGCTCCGTGCTGAAGTT	1760
Db	541	GAAAAAGTCTAGTCCCAAGATGATGCTCA	GAGTCTGTGCTCCGTGCTGAAGTT	600
Qy	1761	TGATGATTAATCTCAACAGAACTTGTCTAC	CCCTCATCATGAAATTCAGAGGAA	1820
Db	601	TGATGATTAATCTCAACAGAACTTGTCTAC	CCCTCATCATGAAATTCAGAGGAA	660
Qy	1821	GACCAACCAACTGATTTGAGATGAGTGGCG	ATATGACACAGAGGAAGATTTTACTGTGTT	1880
Db	661	GACCAACCAACTGATTTGAGATGAGTGGCG	ATATGACACAGAGGAAGATTTTACTGTGTT	720
Qy	1881	GCAGCGCTTCTTGAAGAGTGAACATGCA	AAAGACTCTGGAGGATTCGACAACTC	1940
Db	721	GCAGCGCTTCTTGAAGAGTGAACATGCA	AAAGACTCTGGAGGATTCGACAACTC	780
Qy	1941	TTTCTTGCTCTGACTGTTTCCACTTCA	GACAGCAAGTCTCACTCCGAGCAGCAGTGC	2000
Db	781	TTTCTTGCTCTGACTGTTTCCACTTCA	GACAGCAAGTCTCACTCCGAGCAGCAGTGC	840
Qy	2001	TCTCTGGAACAATATGCTGAGGCTGTG	GCCAGAAAGACAGCTGTCATTAAGTTGAAA	2060
Db	841	TCTCTGGAACAATATGCTGAGGCTGTG	GCCAGAAAGACAGCTGTCATTAAGTTGAAA	900

QY 2061 CAAGATCAATATCATCATGTCGGAACGAGTCCAGCCGTTTCTGAGAGCCTTACAGAACAG 2120
 Db 901 CAAGATCAATATCATCATGTCGGAACGAGTCCAGCCGTTTCTGAGAGCCTTACAGAACAG 960
 QY 2121 CATGCGAGGTCATGAGGACCTGAGCTCCATGAGGAGGAGAGAGCCCTTCTGCTTGACCC 2180
 Db 961 CATGCGAGGTCATGAGGACCTGAGCTCCATGAGGAGGAGAGAGCCCTTCTGCTTGACCC 1020
 QY 2181 TACCTGAGTGCATGAGCCTGAGACCTGAGACATCAAGTTGTGCTGCTGAGGAGATTC 2240
 Db 1021 TACCTGAGTGCATGAGCCTGAGACCTGAGACATCAAGTTGTGCTGCTGAGGAGATTC 1080
 QY 2241 TCTGACCCGTCGCAATGGAATTTGGCTCCAAACAGAGACCTCCCGATGTACACACACA 2300
 Db 1081 TCTGACCCGTCGCAATGGAATTTGGCTCCAAACAGAGACCTCCCGATGTACACACACA 1140
 QY 2301 GTATCGGGGACTGTCTATCATAGTGAGG 2327
 Db 1141 GTATCGGGGACTGTCTATCATAGTGAGG 1167

RESULT 11

ABA08297 standard; cDNA; 802 BP.

ABA08297;

11-JAN-2002 (first entry)

Human phospholipase B homologue-encoding cDNA, SEQ ID NO:73.

Human; cytokine; cell proliferation; cell differentiation; cell growth factor;
 hemotopoiesis regulation; tissue growth; immunomodulation; activin;
 inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 proliferation; metastasis; cancer; tumor; haematopoietic disorder;
 myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 chronic inflammatory condition; proliferative retinopathy;
 atherosclerosis; coronary heart disease; arterial ischaemia;
 bone disorder; osteoporosis; vascular growth disorder;
 tissue regeneration; wound healing; infection; immune disorder;
 cell culture; drug screening; gene therapy; antiinflammatory;
 antistatic; antiahrthritic; haemostatic; antiarteriosclerotic;
 cyostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 antifungal; vulnery; antitumor; ss.

Homo sapiens.

MO200157188-A2.

09-AUG-2001.

05-FEB-2001; 2001WO-US03800.

03-FEB-2000; 2000US-0496914.

27-APR-2000; 2000US-0560875.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Drmanac RT;

WPI; 2001-457740/49.

P-PDB; ABB11053.

Human proteins and DNA encoding sequences useful for preventing,
 treating or ameliorating a medical condition in a mammalian subject
 e.g. arthritis and cancer -

Claim 1; Page 365; 1963pp; English.

Sequences ABB10981-ABB12310 represent 1350 novel human polypeptides, and
 sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 invention also relates to vectors and recombinant host cells comprising a

CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC hemotopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation and metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a cDNA encoding a
 CC novel human polypeptide of the invention.

XX SQ Sequence 802 BP; 206 A; 236 C; 211 G; 149 T; 0 other;

Query Match 15.2%; Score 667.4; DB 22; Length 802;

Best Local Similarity 99.9%; Pred. No. 3.5e-177; Index 0; Gaps 0;

Matches 668; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3057 GCTTGAACCACTTGAAGCAAAACAGAGACCTCTGAGCAGAGATGCCATCAC 3116
 Db 134 GCTTGAACCACTTGAAGCAAAACAGAGACCTCTGAGCAGAGATGCCATCAC 193
 QY 3117 CTGTCCCACTCAGATGAGCCCTTCTGAGAACCCCTCGAATATGTAATACATGCC 3176
 Db 194 CTGTCCCACTCAGATGAGCCCTTCTGAGAACCCCTCGAATATGTAATACATGCC 253
 QY 3177 CATCAAGCCAGCCATTGAGAACTGGGGCAGTGACTTCTGTTCACAGATGGAAGCTTC 3236
 Db 254 CATCAAGCCAGCCATTGAGAACTGGGGCAGTGACTTCTGTTCACAGATGGAAGCTTC 313
 QY 3237 CAATAGTGTTCACCACTCTGTCCACCAAGCTCCAGCAGACATCAAAATGTGAGCGC 3296
 Db 314 CAATAGTGTTCACCACTCTGTCCACCAAGCTCCAGCAGACATCAAAATGTGAGCGC 373
 QY 3297 CTTGGGTGACTCTGTGACTACAGCAGTGGAGCTGACCAAAACACTCCAGTACCTACC 3356
 Db 374 CTTGGGTGACTCTGTGACTACAGCAGTGGAGCTGACCAAAACACTCCAGTACCTACC 433
 QY 3357 CAATCTTGAAGAGGAGCTCTCTTGGAGCAATGAGAGGAGTGGAACTTGAAGCTACAC 3416
 Db 434 CAATCTTGAAGAGGAGCTCTCTTGGAGCAATGAGAGGAGTGGAACTTGAAGCTACAC 493
 QY 3417 CACACTGCCCAACATCTGAGAGAGTTCACCTTACTCTCTTGGCTCTCTACAGAC 3476
 Db 494 CACACTGCCCAACATCTGAGAGAGTTCACCTTACTCTCTTGGCTCTCTACAGAC 553
 QY 3477 CTGGGAGGGAGACAGCACTAATGTGCGAGCGAGGCGCCAGACTGAGGACATGCC 3536
 Db 554 CTGGGAGGGAGACAGCACTAATGTGCGAGCGAGGCGCCAGACTGAGGACATGCC 613

QY 3537 AGCCAGGCTGGAGCTGTAGAGCGAATGAAAAAGCCCGACATCACTGAGAA 3596
DB 614 AGCCAGGCTGGAGCTGTAGAGCGAATGAAAAAGCCCGACATCACTGAGAA 673
QY 3597 AGACTGGAAGCTGTGTCACTCTTCAATTTGGGTCACGACTTGTGTCACTGAGAA 3656
DB 674 AGACTGGAAGCTGTGTCACTCTTCAATTTGGGTCACGACTTGTGTCACTGAGAA 733
QY 3657 TCCGAGAGCCCACTTGGCCAGCAATATGTTCACGACATCCAAAGAGCCCTGGACATCCT 3716
DB 734 TCCGAGAGCCCACTTGGCCAGCAATATGTTCACGACATCCAAAGAGCCCTGGACATCCT 793
QY 3717 CTCTGAGGA 3725
DB 794 CTCTGAGGA 802

RESULT 12
AAH9765
ID AAH9765 standard; cDNA; 802 BP.
XX
AC AAH9765;
XX
DT 16-OCT-2001 (first entry)
XX
DE Human protein encoding cDNA sequence SEQ ID NO:600.
XX
KM Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KM anti-inflammatory; antirheumatic; antiarthritic; immunosuppressive;
KM antibacterial; endocrine; cardiac; central nervous system; vitruide;
KM anti-HIV; fungicide; antimutagen; cardiovascular; antianemic; anaemia;
KM antiagregant; haemostatic; vulnary; antileuk; osteoporosis; eczema;
KM dermatological; antiallergic; antiaesthetic; antidiabetic; cytostatic;
KM neuroprotective; antiparkinsonian; infection;
KM immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KM antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KM cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KM genetic disease; haematopoietic disorder; platelet disorder; asthma;
KM thrombocytopenia; osteoporosis; severe combined immunodeficiency;
KM allergic rhinitis; diabetes; multiple sclerosis; depression;
KM Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KM neurological disorder; ss.
XX
OS Homo sapiens.
XX
PN W0200153455-A2.
XX
PD 26-JUL-2001.
XX
PF 22-DEC-2000; 2000WO-US35017.
XX
PR 23-DEC-1999; 99US-0471275.
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-457603/49.
XX
P-PSDB; AAM25824.
XX
PT Isolated human polynucleotides encoding polypeptides, useful for the
XX
XX treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
XX
XX Claim 1; Page 634; 1217p; English.
XX
XX AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
XX
XX AAM25863. The proteins can have activities based on the tissues and
XX
XX cells they are expressed in, such as: anti-inflammatory; antirheumatic;
XX
XX antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;
XX
XX central nervous system; vitruide; anti-HIV; fungicide; antimutagen;
XX
XX cardiovascular; antianemic; antiagregant; haemostatic; vulnary;

CC antiulcer; osteopathic; dermatological; antiallergic; antiaesthetic;
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nocotropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and
CC neurological disorders.
XX
SQ Sequence 802 BP; 206 A; 236 C; 211 G; 149 T; 0 other;
Query Match 15.2%; Score 667.4; DB 22; Length 802;
Best Local Similarity 99.9%; Pred. No. 3,5e-177;
Matches 668; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3057 GCTTGAACCACTTGGAGCAAAACAGAGACCTTGAACCTTGAAGCAGAGATGCCATCAC 3116
DB 134 GCTTGAACCACTTGGAGCAAAACAGAGACCTTGAACCTTGAAGCAGAGATGCCATCAC 193
QY 3117 CTGTCCCACTGAAATGAGCCCTTCTGAGAACCCCTCGGAATATGTAACATCAACATCC 3176
DB 194 CTGTCCCACTGAAATGAGCCCTTCTGAGAACCCCTCGGAATATGTAACATCAACATCC 253
QY 3177 CATCAAGCCAGCATTGAGAGAGTGGGAGTACTTCTGTGTACAGAGTGAAGGCTTC 3236
DB 254 CATCAAGCCAGCATTGAGAGAGTGGGAGTACTTCTGTGTACAGAGTGAAGGCTTC 313
QY 3237 CAATAGTGTTCACACTTGTGTCCACAGCTCCGACAGACAGATCAATGAATGTGGCGC 3296
DB 314 CAATAGTGTTCACACTTGTGTCCACAGCTCCGACAGACAGATCAATGAATGTGGCGC 373
QY 3297 CTTGGGTGATCTCTGTACATACAGAGTGGAGTTCACCAAACTCCAGTACCTAAC 3356
DB 374 CTTGGGTGATCTCTGTACATACAGAGTGGAGTTCACCAAACTCCAGTACCTAAC 433
QY 3357 CACATCTTGGAGGGGACTCTTGGAGCATTGGAGGGGATGGGAATTGGAGACTACAC 3416
DB 434 CACATCTTGGAGGGGACTCTTGGAGCATTGGAGGGGATGGGAATTGGAGACTACAC 493
QY 3417 CACATCTTGGAGGGGACTCTTGGAGCATTGGAGGGGATGGGAATTGGAGACTACAC 3476
DB 494 CACATCTTGGAGGGGACTCTTGGAGCATTGGAGGGGATGGGAATTGGAGACTACAC 553
QY 3477 CTGGAGAGGAGACAGCAGACTAAATGTGGAGCGGAGGGGCGAGAGCTAGGACATGCC 3536
DB 554 CTGGAGAGGAGACAGCAGACTAAATGTGGAGCGGAGGGGCGAGAGCTAGGACATGCC 613
QY 3537 AGCCAGAGCCCTGGAGCCTGTGTAGAGCAATGAAAAAGCCCGACATCACTGAGAA 3596
DB 614 AGCCAGAGCCCTGGAGCCTGTGTAGAGCAATGAAAAAGCCCGACATCACTGAGAA 673
QY 3597 AGACTGGAAGCTGTGTCACTCTTCAATTTGGGTCACGACTTGTGTCACTGAGAA 3656
DB 674 AGACTGGAAGCTGTGTCACTCTTCAATTTGGGTCACGACTTGTGTCACTGAGAA 733
QY 3657 TCCGAGAGCCCACTTGGCCAGCAATATGTTCACGACATCCAAAGAGCCCTGGACATCCT 3716
DB 734 TCCGAGAGCCCACTTGGCCAGCAATATGTTCACGACATCCAAAGAGCCCTGGACATCCT 793
QY 3717 CTCTGAGGA 3725
DB 794 CTCTGAGGA 802

RESULT 13
AAD37412

ID AAD37412 standard; DNA; 608 BP.
 XX AAD37412;
 XX
 DT 27-AUG-2002 (first entry)
 DE Human phospholipase-like enzyme encoding DNA #2.
 XX
 KM Human, phospholipase-like enzyme; cancer; inflammation; Pick's disease;
 KM cardiovascular disorder; central nervous system disorder; brain injury;
 KM chronic obstructive pulmonary disease; cerebrovascular disease; dementia;
 KM Alzheimer's disease; Parkinson's disease; corticobasal degeneration;
 KM motor neuron disease; Huntington's disease; Creutzfeldt Jacob dementia;
 KM schizophrenia; Korsakoff's psychosis; pain; epilepsy; multiple sclerosis;
 KM sciatica; stroke; age associated memory impairment; allergy; asthma;
 KM allergic rhinitis; hay fever; atopic dermatitis; cardiovascular disease;
 KM anaphylaxis; inflammation; acute respiratory distress syndrome; diabetes;
 KM chronic obstructive pulmonary disease; emphysema; obesity; anorexia;
 KM overweight; cachexia; bulimia; hypertension; coronary artery disease;
 KM type-II diabetes; hyperlipidaemia; gall bladder disease; osteoarthritis;
 KM gout; sleep apnoea; respiratory problem; polycystic ovarian syndrome;
 KM thrombolytic disease; reduced fertility; pregnancy; stress incontinence;
 KM hirsutism; menstrual irregularity; depression; enzyme; ds.
 KM
 OS Homo sapiens.
 XX
 PN WO200231161-A2.
 XX
 PD 18-APR-2002.
 XX
 PF 09-OCT-2001; 2001WO-EP11641.
 XX
 PR 10-OCT-2000; 2000US-238445P.
 XX
 PR 26-DEC-2000; 2000US-257293P.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Zhu Z;
 XX
 DR WP1; 2002-426287/45.
 XX
 PT New human phospholipase-like enzyme polypeptide useful for screening
 PT agents, and in the treatment of cancer, inflammation, diabetes,
 PT obesity, a central nervous system disorder, or a cardiovascular
 PT disorder.
 XX
 PS Disclosure; Fig 5; 144pp; English.
 XX
 CC The present invention relates to novel human phospholipase-like enzymes
 CC and polynucleotides encoding such proteins. Sequences of the invention
 CC are useful for producing a medicament for modulating the activity of
 CC phospholipase in a disease such as cancer, inflammation, cardiovascular
 CC disorders, chronic obstructive pulmonary diseases, central nervous system
 CC (CNS) disorders such as brain injuries, cerebrovascular disease, dementia
 CC (Alzheimer's disease), Parkinson's disease, corticobasal degeneration,
 CC motor neuron disease, Huntington's disease, Creutzfeldt
 CC Jacob dementia, schizophrenia with dementia, Korsakoff's psychosis,
 CC pain associated with CNS (e.g. epilepsy, failed back surgery syndrome,
 CC sciatica), multiple sclerosis, stroke, age associated memory impairment,
 CC allergic disease including asthma, allergic rhinitis (hay fever), atopic
 CC dermatitis, anaphylaxis and inflammation, cardiovascular disease, chronic
 CC obstructive pulmonary disease, acute respiratory distress syndrome, gout,
 CC diabetes, emphysema or obesity. They are also used for treating anorexia,
 CC overweight, cachexia, bulimia, hypertension, type-II diabetes, coronary
 CC artery disease, hyperlipidaemia, gall bladder disease, osteoarthritis,
 CC sleep apnoea and respiratory problems, cancer (e.g. breast, prostate,
 CC colon cancer), thrombolytic disease, reduced fertility, polycystic
 CC ovarian syndrome, complications of pregnancy, menstrual irregularity,
 CC hirsutism, stress incontinence and depression. The present sequence is
 CC human phospholipase like enzyme encoding DNA.
 XX
 SQ Sequence 608 BP; 140 A; 187 C; 158 G; 123 T; 0 other;

Query Match 13.2%; Score 576.2; DB 24; Length 608;
 Best Local Similarity 97.8%; Pred. No. 1,5e-151;
 Matches 584; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 2654 TGGACGCTCTGCATAGAGAGGTGCCAGAGTCTGTGCTCAACTGTGAGACTTCTGAAACC 2713
 DB 11 TTGATGGGCTGCATAGAGAGGTGCCAGAGTCTGTGCTCAACTGTGAGACTTCTGAAACC 70
 QY 2714 CCACATATCATGGCGGAGGTGTTCTCTGGGAAACCCAGACAAAGTCCAGTCCAGAGGCCA 2773
 DB 71 CCACATATCATGGCGGAGGTGTTCTCTGGGAAACCCAGACAAAGTCCAGTCCAGAGGCCA 130
 QY 2774 GCGTTTGTGTACTGCTGCTTCTGACCTCGGGGAGAACTCCCAAGAGCTTAGCCAGCTGG 2833
 DB 131 GCGTTTGTGTACTGCTGCTTCTGACCTCGGGGAGAACTCCCAAGAGCTTAGCCAGCTGG 190
 QY 2834 AGGCTTTAGCCGAGAGCTTACCCGAGAGCATATGCGAGCTGTGGGGTCAAGCGCTTAG 2893
 DB 191 AGGCTTTAGCCGAGAGCTTACCCGAGAGCATATGCGAGCTGTGGGGTCAAGCGCTTAG 250
 QY 2894 ACAAGGAGAGAGAGCTTCTGCTGTGTGCTGAGAGCCCTTTCCAGAACATCCAGCTCCCG 2953
 DB 251 ACAAGGAGAGAGAGCTTCTGCTGTGTGCTGAGAGCCCTTTCCAGAACATCCAGCTCCCG 310
 QY 2954 TCCTGGCGATGGGCTCCAGATACGTCCTCTTTTCCCGAGACTGATCCACCAATC 3013
 DB 311 TCCTGGCGATGGGCTCCAGATACGTCCTCTTTTCCCGAGACTGATCCACCAATC 370
 QY 3014 AGAATTTCCATCTCCAGCTGGGCCAGAGCCCTTTGACCAATATGCTTGAACCACTTGGAA 3073
 DB 371 AGAATTTCCATCTCCAGCTGGGCCAGAGCCCTTTGACCAATATGCTTGAACCACTTGGAA 430
 QY 3074 GCAAAACAGAGACCTCTGAGACCTGAGAGAGAGATGCGCATACCTCTCCACATCAATG 3133
 DB 431 GCAAAACAGAGACCTCTGAGACCTGAGAGAGATGCGCATACCTCTCCACATCAATG 490
 QY 3134 AGCCCTTCTGAGAAACCTCTGGAATAGTACTACAGTACCCATCAAGCCAGCATTTG 3193
 DB 491 AGCCCTTCTGAGAAACCTCTGGAATAGTACTACAGTACCCATCAAGCCAGCATTTG 550
 QY 3194 AGAAGTGGGCGAGTGAATCTCTGTGTACAGAGTGAAGGCTTCCAAATAGTTTCAA 3250
 DB 551 AGAAGTGGGCGAGTGAATCTCTGTGTACAGAGTGAAGGCTTCCAAATAGTTTCAA 607

RESULT 14
 AAL01390
 ID AAL01390 standard; cDNA; 572 BP.
 XX
 AC AAL01390;
 XX
 DT 21-NOV-2001 (first entry)
 XX
 DE Human reproductive system related antigen cDNA SEQ ID NO: 1391.
 XX
 KM Human; reproductive system related antigen; reproductive system disorder;
 KM cancer; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200155320-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01339.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 XX
 PR 04-FEB-2000; 2000US-0180628.
 XX
 PR 24-FEB-2000; 2000US-0184664.
 XX
 PR 02-MAR-2000; 2000US-0186350.
 XX
 PR 16-MAR-2000; 2000US-0189874.
 XX
 PR 17-MAR-2000; 2000US-0190076.
 XX
 PR 18-APR-2000; 2000US-0198123.

PR	20-OCT-2000,	2000US-0241221.
PR	20-OCT-2000,	2000US-0241785.
PR	20-OCT-2000,	2000US-0241786.
PR	20-OCT-2000,	2000US-0241787.
PR	20-OCT-2000,	2000US-0241808.
PR	20-OCT-2000,	2000US-0241809.
PR	20-OCT-2000,	2000US-0241826.
PR	01-NOV-2000,	2000US-0244611.
PR	08-NOV-2000,	2000US-0246474.
PR	08-NOV-2000,	2000US-0246475.
PR	08-NOV-2000,	2000US-0246476.
PR	08-NOV-2000,	2000US-0246477.
PR	08-NOV-2000,	2000US-0246478.
PR	08-NOV-2000,	2000US-0246523.
PR	08-NOV-2000,	2000US-0246524.
PR	08-NOV-2000,	2000US-0246525.
PR	08-NOV-2000,	2000US-0246526.
PR	08-NOV-2000,	2000US-0246527.
PR	08-NOV-2000,	2000US-0246528.
PR	08-NOV-2000,	2000US-0246532.
PR	08-NOV-2000,	2000US-0246609.
PR	08-NOV-2000,	2000US-0246610.
PR	08-NOV-2000,	2000US-0246611.
PR	08-NOV-2000,	2000US-0246613.
PR	17-NOV-2000,	2000US-0249207.
PR	17-NOV-2000,	2000US-0249208.
PR	17-NOV-2000,	2000US-0249209.
PR	17-NOV-2000,	2000US-0249210.
PR	17-NOV-2000,	2000US-0249211.
PR	17-NOV-2000,	2000US-0249212.
PR	17-NOV-2000,	2000US-0249213.
PR	17-NOV-2000,	2000US-0249214.
PR	17-NOV-2000,	2000US-0249215.
PR	17-NOV-2000,	2000US-0249216.
PR	17-NOV-2000,	2000US-0249217.
PR	17-NOV-2000,	2000US-0249218.
PR	17-NOV-2000,	2000US-0249244.
PR	17-NOV-2000,	2000US-0249245.
PR	17-NOV-2000,	2000US-0249264.
PR	17-NOV-2000,	2000US-0249265.
PR	17-NOV-2000,	2000US-0249297.
PR	17-NOV-2000,	2000US-0249299.
PR	17-NOV-2000,	2000US-0249300.
PR	01-DEC-2000,	2000US-0250160.
PR	01-DEC-2000,	2000US-0250391.
PR	05-DEC-2000,	2000US-0251030.
PR	05-DEC-2000,	2000US-0251988.
PR	05-DEC-2000,	2000US-0256718.
PR	06-DEC-2000,	2000US-0251479.
PR	08-DEC-2000,	2000US-0251856.
PR	08-DEC-2000,	2000US-0251866.
PR	08-DEC-2000,	2000US-0251869.
PR	08-DEC-2000,	2000US-0251989.
DR	08-DEC-2000,	2000US-0251990.
PR	11-DEC-2000,	2000US-0254097.
PR	05-JAN-2001,	2001US-02589678.
XX	(HUMA-) HUMAN GENOME SCI INC.	
PA	Rosen CA, Barash SC, Ruben SM,	
XX	WPI, 2001-465570/50.	
XX	P-PsDB; AAM95420.	
DR		
XX		
PT	Isolated nucleic acid molecule encoding a reproductive system antigen	
CC	is used in preventing, treating or ameliorating a medical condition -	
XX	Claim 1; SEQ ID NO 1391, 1297bp + Sequence Listing; English.	
XX	The present invention provides the protein and coding sequences of a	
CC	number of human reproductive system related antigens. These can be used	
CC	in the prevention and treatment of reproductive system disorders,	
CC	including cancer. The present sequence is a coding sequence of the	

CC invention.
XX Sequence 572 BP; 146 A; 159 C; 157 G; 106 T; 4 other;
SQ Query Match 9.4%; Score 411.2; DB 22; Length 572;
Best Local Similarity 96.7%; Pred. No. 4,6e-105;
Matches 492; Conservative 2; Mismatches 6; Indels 9; Gaps 7;
QY 3042 CCTTTGGACCAATATGCTTTGAACCACTTGGAAAGCAAGACCCCTGGACCTGGAGAGC 3101
DB 15 CCTTTGAGCCAAATATGCTTTGAACCACTTGGAAAGCAAGACCCCTGGAGAGC 74
QY 3102 AGAGATGCCCATCACTGTGTCCCACTAGATAGAGCCCTTCTTGAAACCCCTCGAATAG 3161
DB 75 AGAGATGCCCATCACTGTGTCCCACTAGATAGAGCCCTTCTTGAAACCCCTCGAATAG 134
QY 3162 TAACTACACGTAACCCCATCAAGCCATTTGAATAGAGAGCTGGGGCAGTGACTTCTGTGAC 3221
DB 135 TAACTACACGTAACCCCATCAAGCCATTTGAATAGAGAGCTGGGGCAGTGACTTCTGTGAC 194
QY 3222 AGAGTGAAGGCTTCCATATAGTTTCCAACTGTCCACAGCTCCGAGCCAGAGACAT 3281
DB 195 AGAGTGAAGGCTTCCATATAGTTTCCAACTGTCCACAGCTCCGAGCCAGAGACAT 254
QY 3282 CAAAGTGATGCGCCCTGTGGTGACTCTGTGACTACAGAGTGGAGCTCGACCAACAA 3341
DB 255 CAAAGTGATGCGCCCTGTGGTGACTCTGTGACTACAGAGTGGAGCTCGACCAACAA 314
QY 3342 CTCAGTGTG-ACTTACCAATCTTTGAGGGGAGCTCTCTT-GGACATTTGAGGGGATGGG 3399
DB 315 CTCAGTGTGACTTACCAATCTTTGAGGGGAGCTCTTGTGGAGCATTTGAGGGGATGGG 374
QY 3400 AA--CTTGGAGACTCACACCACTGCCCCAATTTGAAATTTCAACCTTAACTCC 3457
DB 375 NAACTTGGAGACTCACACCACTGCCCCAATTTGAAATTTCAACCTTAACTCC 434
QY 3458 TTGGCTTCTCT-ACCAGCACCCTGGG--AGGGAGACAGCAGACT-AAATGTGGCAGCGAA 3513
DB 435 TTGGCTTCTCTGACACAGACCTGGGGAGGGGAGACAGCAGACTGAATGTGGCAGCGAA 494
QY 3514 GGGGCCAGAGCTA-GGACATGCGAGCCC 3541
DB 495 GGGGCCAGAGTTAGGGGAGCAGCGAGCCC 523
RESULT 15
ABL96843
ID ABL96843 standard; cDNA; 572 BP.
XX ABL96843;
AC
XX
DT 21-JUN-2002 (first entry)
XX
XX Human testicular antigen encoding cDNA SEQ ID NO: 511.
DE
XX
XX Human; testicular antigen; testes; cancer; metastasis; immune disorder;
KW reproductive system disorder; urinary system disorder; gene therapy;
KW cardiovascular disorder; respiratory disorder; neurological disorder;
KW gastrointestinal disease; infection; cyostatic; gene; ss.
XX
XX Homo sapiens.
OS
XX
XX W0200155317-A2.
PN
XX
XX 02-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US01329.
PF
XX
XX 31-JAN-2000; 2000US-0179065.
PR
XX 04-FEB-2000; 2000US-0180628.
PR
XX 24-FEB-2000; 2000US-0184664.
PR
XX 02-MAR-2000; 2000US-0186350.
PR
XX 16-MAR-2000; 2000US-0189874.
PR

PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0218230.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225477.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226581.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 27-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR

PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

XX (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Barash SC, Ruben SM;

XX DR WPI; 2001-483232/52.

XX PT Nucleic acids encoding 973 human testicular antigen polypeptides,
XX useful for preventing, diagnosing and/or treating testicular cancer -

XX PS Claim 1; SEQ ID NO 511; 766bp; English.

XX CC The present invention provides the protein and coding sequences of 973
XX human testicular antigens, and fragments of their genomic sequences. The
XX sequences can be used in the treatment of cardiovascular, urinary system,

CC reproductive system, immune, respiratory, neurological and
CC gastrointestinal disorders, infections, and particularly cancer,
CC especially testicular cancers. The present sequence is a cDNA of the
CC invention.

XX SQ Sequence 572 BP; 146 A; 159 C; 157 G; 106 T; 4 other;

Query Match 9.4%; Score 411.2; DB 23; Length 572;
Best Local Similarity 96.7%; Pred. No. 4,6e-105;
Matches 492; Conservative 2; Mismatches 6; Indels 9; Gaps 7;

QY 3042 CCTTGGACCAATATGCTTGAACCACTTGGAAACAAAGAGACCTTGAGCTGAGAGC 3101
DB 15 CTTTGGACCAATATGCTTGAACCACTTGGAAACAAAGAGACCTTGAGCTGAGAGC 74
QY 3102 AGAGATGCCCATACCTGTGCCCTCAGATGAGCCCTTCTTGAAGAACCTTGGAATAG 3161
DB 75 AGAGATGCCCATACCTGTGCCCTCAGATGAGCCCTTCTTGAAGAACCTTGGAATAG 134
QY 3162 TAACTACAGTACCCCATCAAGCCAGCCATTTGAACTGGGGGAGTACTTCTGTGTAC 3221
DB 135 TAACTACAGTACCCCATCAAGCCAGCCATTTGAACTGGGGGAGTACTTCTGTGTAC 194
QY 3222 AGAGTGAAGGCTTCCATATGCTTCCAACTGTCCACAGCTCCGACAGACAT 3281
DB 195 AGAGTGAAGGCTTCCATATGCTTCCAACTGTCCACAGCTCCGACAGACAT 254
QY 3282 CAAAGTGTGCGCCCTGGGTGATCTCTGATACAGCGTGGAGCTGACCAACAA 3341
DB 255 CAAAGTGTGCGCCCTGGGTGATCTCTGATACAGCGTGGAGCTGACCAACAA 314
QY 3342 CTCGAGTG-ACTTACCACATCTTGGAGGGGACTCTCTT-GGAGCATTTGAGGGGATGG 3399
DB 315 CTCGAGTGAACTTACCACATCTTGGAGGGGACTCTTGGAGCATTTGAGGGGATGG 374
QY 3400 AA--CTTGAAGACTACACCACTGCCCCAATTTGAAAGATTCAACCTTACCTCC 3457
DB 375 NAACTGGAGACTCAACACACTGCCCAATCTTGAAGATTCAACCTTACCTCC 434
QY 3458 TTGGCTTCTCT-ACGACCACTGGG--AGGGGACAGAGACT-AAATGTGGACGGGAA 3513
DB 435 TTGGCTTCTCTGACCAAGCACTTGGGAGGGGGACAGAGACTGAATGTGGACGGGAA 494
QY 3514 GGGGCCAGAGCTA-GGAGCATGCGAGCCC 3541
DB 495 GGGGCCAGAGTTAGGGGACATGCCAGCCC 523

Search completed: January 6, 2004, 08:00:25
Job time : 1748 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 6, 2004, 04:33:27 ; Search time 15537 Seconds

(without alignments)
11524.840 Million cell updates/sec

Title: US-10-054-691-1

Perfect score: 4377
Sequence: 1 atcgggcgctgcgcagcat.....tgcacactgtggccctctag 4377

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

GenBank: 1: gb_ba: 2: gb_hcg: 3: gb_in: 4: gb_cm: 5: gb_ov: 6: gb_pat: 7: gb_ph: 8: gb_pl: 9: gb_dr: 10: gb_ro: 11: gb_sts: 12: gb_sy: 13: gb_un: 14: gb_vl: 15: em_ba: 16: em_fun: 17: em_hum: 18: em_mu: 19: em_mu: 20: em_om: 21: em_ov: 22: em_ov: 23: em_pat: 24: em_ph: 25: em_pl: 26: em_ro: 27: em_sts: 28: em_un: 29: em_vl: 30: em_hcg_hum: 31: em_hcg_inv: 32: em_hcg_other: 33: em_hcg_mus: 34: em_hcg_pln: 35: em_hcg_rod: 36: em_hcg_mam: 37: em_hcg_vrt: 38: em_sy: 39: em_hcgo_hum: 40: em_hcgo_mus: 41: em_hcgo_other:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4377	100.0	4377	6	AX492941
2	3461.2	79.1	3648	6	AX417817
3	3124	71.4	4612	4	OCPRHLP
4	2962	67.7	4608	10	AF045454
5	2687	61.4	4590	10	D63648
6	2687	61.4	4513	6	E13935
7	1362	31.1	1835	6	AX171962
8	1319.4	30.1	1790	9	BC042674
9	970	22.2	2477	6	AX173706
10	970	22.2	2477	6	AK055428
11	576.2	13.2	608	6	AX417821
12	372.2	8.5	583	6	AX417820
13	322.2	7.4	820	10	BC033606
14	251	5.7	265	6	AX417822
15	237.4	5.4	446	6	AX417823
16	205	4.7	43543	6	AX171964
17	205	4.7	180465	9	AC074011
18	177.2	4.0	163595	2	AC022076
19	120.6	2.8	1672	3	AY069174
20	109	2.5	103926	9	AC093164
21	107.8	2.5	357	6	AX14612
22	107.2	2.4	50250	2	AC017341
23	107.2	2.4	179283	3	AC009383
24	107.2	2.4	299903	3	AB003514
25	94.6	2.2	101340	2	AC018104
26	94.6	2.2	190668	3	AC007419
27	94.6	2.2	260731	3	AB003611
28	85.6	2.0	217253	2	AC102372
29	85.6	1.9	214038	2	AC141490
30	81.6	1.9	214038	2	AC141490
31	75.6	1.7	217253	2	AC102372
32	70.4	1.6	134246	9	AC125617
33	68.6	1.6	684	4	CFA243475
34	68.2	1.6	125020	9	AF429315
35	63.2	1.4	125020	9	AF429315
36	61.6	1.4	631	11	BV055548
37	59.8	1.4	44237	2	AC024847
38	59.8	1.4	267118	2	AC006889
39	58	1.3	198736	5	AL929023
40	52.6	1.2	2000	6	AX655393
41	52.6	1.2	123789	2	AC110674
42	52.2	1.2	2000	6	AX655393
43	49.2	1.1	232304	2	AC111030
44	47.8	1.1	34305	3	CE002812
45	47.2	1.1	212127	5	BX248515

ALIGNMENTS

RESULT 1	AX492941	4377 bp	DNA	linear	PAT 26-SEP-2002
LOCUS	AX492941				
DEFINITION	Sequence 1 from Patent WO02059328.				
ACCESSION	AX492941				
VERSION	AX492941.1	GI:23338611			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens (human)				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1				
AUTHORS	Turner Jr,C.A., Miranda,M. and Yu,X.				
TITLE	Human lipase and polynucleotides encoding the same				
JOURNAL	Patent: WO 02059328-A 1 01-AUG-2002;				

LEXICON GENETICS INC (US)
Location/Qualifiers
1. 4377
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 1085 a 1217 c 1155 g 920 t
ORIGIN

Query Match 100.0%; Score 4377; DB 6; Length 4377;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGGGTGGGGCCAGGCAATTTTCTCTGAGAGTGGTGGTCTTCTGGGGCAAGGAGC 60
Db 1 ATGGGGGTGGGGCCAGGCAATTTTCTCTGAGAGTGGTGGTCTTCTGGGGCAAGGAGC 60

QY 61 CCTCAGATCCATCTCTCTAGAAAGTACATTTGAGAGGGCAGTATGGCCAGAGACC 120
Db 61 CCTCAGATCCATCTCTCTAGAAAGTACATTTGAGAGGGCAGTATGGCCAGAGACC 120

QY 121 CTGAAGAATTTCTCATTTCCATGCAACCAATTAATAGAGTGAATATGCTTTCTTAA 180
Db 121 CTGAAGAATTTCTCATTTCCATGCAACCAATTAATAGAGTGAATATGCTTTCTTAA 180

QY 181 TCAGTTCACTCTCTGAAGCCTTCTGATATTAATTTGTCAGCCATTTGGCAATCTGAA 240
Db 181 TCAGTTCACTCTCTGAAGCCTTCTGATATTAATTTGTCAGCCATTTGGCAATCTGAA 240

QY 241 ATTCTCTCAGACCCAGGAGCGGGGATCTGAGAGAGCAAGCTGACTGAAGGCCAGC 300
Db 241 ATTCTCTCAGACCCAGGAGCGGGGATCTGAGAGAGCAAGCTGACTGAAGGCCAGC 300

QY 301 CAGGTGCAATGGAGAGTGAACAGTCTTTCAACATCATCAATATTTCAATCTTCT 360
Db 301 CAGGTGCAATGGAGAGTGAACAGTCTTTCAACATCATCAATATTTCAATCTTCT 360

QY 361 GTTCCAAATGCTGTGTGTCACACTGAAAGAGTCAATCCCAAGTGGTGGTGAAGAC 420
Db 361 GTTCCAAATGCTGTGTGTCACACTGAAAGAGTCAATCCCAAGTGGTGGTGAAGAC 420

QY 421 TTGTGATTCAGGCTCAAGACTGCTGTGAGAAACATGAAAGAGCACTGCAACTTT 480
Db 421 TTGTGATTCAGGCTCAAGACTGCTGTGAGAAACATGAAAGAGCACTGCAACTTT 480

QY 481 CAATTTGACTGGAAGCTCATATGTGTCTTCACTGATATGCAAGCCAGTGTACCTGTGC 540
Db 481 CAATTTGACTGGAAGCTCATATGTGTCTTCACTGATATGCAAGCCAGTGTACCTGTGC 540

QY 541 CCTCTGCTCAACAGATGGGCTTGGCGGGCGGCGGTGATGAGTGGGGGTCTG 600
Db 541 CCTCTGCTCAACAGATGGGCTTGGCGGGCGGCGGTGATGAGTGGGGGTCTG 600

QY 601 GACTACCTGAGCAGAGAGTCCCGAGAGCAATTTGTAAACCTGTGTGACCTCTCTGAGTT 660
Db 601 GACTACCTGAGCAGAGAGTCCCGAGAGCAATTTGTAAACCTGTGTGACCTCTCTGAGTT 660

QY 661 GCAGAGTCTCTGCTCAGTATCAACGAGCACTTGGCTCAGCCCTGACCAAGAGCCTGTAAAT 720
Db 661 GCAGAGTCTCTGCTCAGTATCAACGAGCACTTGGCTCAGCCCTGACCAAGAGCCTGTAAAT 720

QY 721 TGCTCAGAGAGACCAACCGGCTGGGCAAGGTGTGATGCAAGGTCTTATCAGAAAGCC 780
Db 721 TGCTCAGAGAGACCAACCGGCTGGGCAAGGTGTGATGCAAGGTCTTATCAGAAAGCC 780

QY 781 TGGAAACGCTCTGCTGCTCAGAGAGTACAGTGAAGAGAGTCTTCAACGCTGTTTC 840
Db 781 TGGAAACGCTCTGCTGCTCAGAGAGTACAGTGAAGAGAGTCTTCAACGCTGTTTC 840

QY 841 CAGGCTTTCTTATGAGACCAACCAATCTTACACTGAGAGACCCCGAATTCAGAGAT 900
Db 841 CAGGCTTTCTTATGAGACCAACCAATCTTACACTGAGAGACCCCGAATTCAGAGAT 900

QY 901 CAGGCTTTCTTATGAGACCAACCAATCTTACACTGAGAGACCCCGAATTCAGAGAT 900
Db 901 CAGGCTTTCTTATGAGACCAACCAATCTTACACTGAGAGACCCCGAATTCAGAGAT 900

QY 901 TCTACACGCTGCGCTGCGATCTCTGAAATAGATGATGAGACCCAGAGAGAAAGAT 960
Db 901 TCTACACGCTGCGCTGCGATCTCTGAAATAGATGATGAGACCCAGAGAGAAAGAT 960

QY 961 GAGCCATTGAGTGAACACCGGAGGCCAATGATGTCCTCTCAGAGAGACCCCTAT 1020
Db 961 GAGCCATTGAGTGAACACCGGAGGCCAATGATGTCCTCTCAGAGAGACCCCTAT 1020

QY 1021 CTGTGAGCTACAGAAACGCACTACCTGACCACTGAGAAACCCCAAGCAAGCTT 1080
Db 1021 CTGTGAGCTACAGAAACGCACTACCTGACCACTGAGAAACCCCAAGCAAGCTT 1080

QY 1081 GAGTAAG 1140
Db 1081 GAGTAAG 1140

QY 1141 ACCTCAGTTCAATGCTGAAAGCCGCTGACATCAAGTAAATGAGAGCCTGGTACTCT 1200
Db 1141 ACCTCAGTTCAATGCTGAAAGCCGCTGACATCAAGTAAATGAGAGCCTGGTACTCT 1200

QY 1201 CTGACGAG 1260
Db 1201 CTGACGAG 1260

QY 1261 TACCGAG 1320
Db 1261 TACCGAG 1320

QY 1321 GCGAAG 1380
Db 1321 GCGAAG 1380

QY 1381 GAAACAG 1440
Db 1381 GAAACAG 1440

QY 1441 CCGTCCAG 1500
Db 1441 CCGTCCAG 1500

QY 1501 GAAAG 1560
Db 1501 GAAAG 1560

QY 1561 GATCTGCTCAATCTCTCCAG 1620
Db 1561 GATCTGCTCAATCTCTCCAG 1620

QY 1621 CTCATGCTGAG 1680
Db 1621 CTCATGCTGAG 1680

QY 1681 CTGAG 1740
Db 1681 CTGAG 1740

QY 1741 TGTCCCTGCTGCTGAG 1800
Db 1741 TGTCCCTGCTGCTGAG 1800

QY 1801 AACAG 1860
Db 1801 AACAG 1860

QY 1861 GAAAGATTTACTGAG 1920
Db 1861 GAAAGATTTACTGAG 1920

QY 1921 GAAAGATTTACTGAG 1980
Db 1921 GAAAGATTTACTGAG 1980

QY 1981 CACTCCGAG 2040
Db 1981 CACTCCGAG 2040

```
Db 1981 GACTCCGAGCACCAGTGTCTCTGGAAACAATATGTGGAGCTGTGGCCAGAAAGACG 2040
Qy 2041 ACTCGTCATTAAGTTTGAAGAAAGATCAATATCAATGTCCGAACCAAGGTCGACGGCTTT 2100
Db 2041 ACTCGTCATTAAGTTTGAAGAAAGATCAATATCAATGTCCGAACCAAGGTCGACGGCTTT 2100
Qy 2101 CTGAGGACCTTACAGAAACAGACATGACAGGCTCATGAGGACTGGCTGCATGACAGGACAGA 2160
Db 2101 CTGAGGACCTTACAGAAACAGACATGACAGGCTCATGAGGACTGGCTGCATGACAGGACAGA 2160
Qy 2161 GCGCCCTTCTGCTTGGCACTTACCTTACATGTCATGCTCCCTGAGACCTTGACATCCAAATT 2220
Db 2161 GCGCCCTTCTGCTTGGCACTTACCTTACATGTCATGCTCCCTGAGACCTTGACATCCAAATT 2220
Qy 2221 GTGGCTGCTTGGGGGATTTCTGACCGCTGGCAATGGAAATGGCTCCAAACCAAGACAGC 2280
Db 2221 GTGGCTGCTTGGGGGATTTCTGACCGCTGGCAATGGAAATGGCTCCAAACCAAGACAGC 2280
Qy 2281 CTCCCGATGTCAACAACAGTATCGGGGACTGTCAATACATGACGAGAGGGGACGGCTCC 2340
Db 2281 CTCCCGATGTCAACAACAGTATCGGGGACTGTCAATACATGACGAGAGGGGACGGCTCC 2340
Qy 2341 CTGGAGAAATGTGACCACTTACCTTAATATCTTGGGAGTTTAAAGAAAACCTCACAGGC 2400
Db 2341 CTGGAGAAATGTGACCACTTACCTTAATATCTTGGGAGTTTAAAGAAAACCTCACAGGC 2400
Qy 2401 TACGCGCTGGGCAACGGGTGATGCGCAATGACAGAAATGCAATCTCTCAATCAAGCTGTTCC 2460
Db 2401 TACGCGCTGGGCAACGGGTGATGCGCAATGACAGAAATGCAATCTCTCAATCAAGCTGTTCC 2460
Qy 2461 GGAGCAAAAGGCTGAGATCTTATGAGCAAGTCCAAACTGTATGACGAAGATGAAGAT 2520
Db 2461 GGAGCAAAAGGCTGAGATCTTATGAGCAAGTCCAAACTGTATGACGAAGATGAAGAT 2520
Qy 2521 GATCATAGATTAATTTCCATGAAAGCTGGAAGTCAATCAAGTGTGATCGAGGACAG 2580
Db 2521 GATCATAGATTAATTTCCATGAAAGCTGGAAGTCAATCAAGTGTGATCGAGGACAG 2580
Qy 2581 GATTTATGTGACACTGACAGATGGAATCTGTATTTGACAGCAACTTTGTTGACAT 2640
Db 2581 GATTTATGTGACACTGACAGATGGAATCTGTATTTGACAGCAACTTTGTTGACAT 2640
Qy 2641 CTCGGAATGCTTGGACGTCCTGCAATAGAGAGTGGCCAGAGTCTTGATCAACTCTG 2700
Db 2641 CTCGGAATGCTTGGACGTCCTGCAATAGAGAGTGGCCAGAGTCTTGATCAACTCTG 2700
Qy 2701 GACTTCCTGAACCCCACTATCATGCGGCAAGTGTCTTGGGAAACCCAGCAAGTCCCA 2760
Db 2701 GACTTCCTGAACCCCACTATCATGCGGCAAGTGTCTTGGGAAACCCAGCAAGTCCCA 2760
Qy 2761 GTGACAGAGGCGCTTTGTGTAACTGCGTTCTGAACCTTGCGGAGAACTTCCAAAG 2820
Db 2761 GTGACAGAGGCGCTTTGTGTAACTGCGTTCTGAACCTTGCGGAGAACTTCCAAAG 2820
Qy 2821 CTAGCCAGGCTGAGGCTTCAAGCCAGCTTACCGGAGAGCAATGGGCAAGCTGTGGGG 2880
Db 2821 CTAGCCAGGCTGAGGCTTCAAGCCAGCTTACCGGAGAGCAATGGGCAAGCTGTGGGG 2880
Qy 2881 TCAGGCGGCTATGACAGCAGAGAGACTTCTGTGGTGTGACGCTTCTTCCAGAAC 2940
Db 2881 TCAGGCGGCTATGACAGCAGAGAGACTTCTGTGGTGTGACGCTTCTTCCAGAAC 2940
Qy 2941 ATTCAGACTCCCTGTCTGCGGAGATGGGCTTCCAGATACGTCCTTGTGGCCCAAGTGC 3000
Db 2941 ATTCAGACTCCCTGTCTGCGGAGATGGGCTTCCAGATACGTCCTTGTGGCCCAAGTGC 3000
Qy 3001 ATCCACCAATGAGAAATTCACATCCGAGCTGGCCAGAGCTTGGACCAATATGCTT 3060
Db 3001 ATCCACCAATGAGAAATTCACATCCGAGCTGGCCAGAGCTTGGACCAATATGCTT 3060
Qy 3061 GAAACAATTGGAAGCAAAACAGAGACCTTGAACCTGAGACAGAGATGCCATCACTGT 3120
Db 3061 GAAACAATTGGAAGCAAAACAGAGACCTTGAACCTGAGACAGAGATGCCATCACTGT 3120
Db 3061 GAAACAATTGGAAGCAAAACAGAGACCTTGAACCTGAGACAGAGATGCCATCACTGT 3120
Qy 3121 CCCACTCAGAAATGAGCCCTTCCAGAACCCCTCGGAATGATCAACATGATCCCATC 3180
Db 3121 CCCACTCAGAAATGAGCCCTTCCAGAACCCCTCGGAATGATCAACATGATCCCATC 3180
Qy 3181 AAGCCAGCATTGAGAACTGGGGCAGTGAATTTCTGTGTACAGAGTGAAGGCTTCAAT 3240
Db 3181 AAGCCAGCATTGAGAACTGGGGCAGTGAATTTCTGTGTACAGAGTGAAGGCTTCAAT 3240
Qy 3241 AGTGTTCAACTCTGTCCACAGCTCCGACAGACATCAAAAGTGTGCGCCCTG 3300
Db 3241 AGTGTTCAACTCTGTCCACAGCTCCGACAGACATCAAAAGTGTGCGCCCTG 3300
Qy 3301 GGTGACTCTGTACATACAGAGTGGAGTTCGACCAACATCCGATGACCTTACCCACA 3360
Db 3301 GGTGACTCTGTACATACAGAGTGGAGTTCGACCAACATCCGATGACCTTACCCACA 3360
Qy 3361 TCTTGGAGGGGACTCTCTTGGAGCAATGGAGGGATGGGAATTGAGACTCACACACA 3420
Db 3361 TCTTGGAGGGGACTCTCTTGGAGCAATGGAGGGATGGGAATTGAGACTCACACACA 3420
Qy 3421 CTGCCCAATCTTGAAGAAATTCAACCTTACTCTTGGCTTCTTACAGACACTG 3480
Db 3421 CTGCCCAATCTTGAAGAAATTCAACCTTACTCTTGGCTTCTTACAGACACTG 3480
Qy 3481 GAGGGGACAGAGCAATTAATGCGAGGGGAGGGGCGAGAGCTTGGGACATGCCAGC 3540
Db 3481 GAGGGGACAGAGCAATTAATGCGAGGGGAGGGGCGAGAGCTTGGGACATGCCAGC 3540
Qy 3541 CAGGCTGGGACCTGTGAGAGCAATGAAGAAACAGCCCGCATCACTTGGAGAAAGAC 3600
Db 3541 CAGGCTGGGACCTGTGAGAGCAATGAAGAAACAGCCCGCATCACTTGGAGAAAGAC 3600
Qy 3601 TGAAGCTGTGACACTTCTTCAATGGGGTCAACGACTTGTGATTACTGTGAGAAATCCG 3660
Db 3601 TGAAGCTGTGACACTTCTTCAATGGGGTCAACGACTTGTGATTACTGTGAGAAATCCG 3660
Qy 3661 GAGGCCCACTTGGCCAGCAATATGTTGACGACATTCACAGGCCCTTGACATCTCTCT 3720
Db 3661 GAGGCCCACTTGGCCAGCAATATGTTGACGACATTCACAGGCCCTTGACATCTCTCT 3720
Qy 3721 GAGAGCTCCCAAGGGCTTTCGTCAACGTTGAGAGTCAATGAGAGTGTGAGCTGATC 3780
Db 3721 GAGAGCTCCCAAGGGCTTTCGTCAACGTTGAGAGTCAATGAGAGTGTGAGCTGATC 3780
Qy 3781 CAGGGCCAAAGGCGGGAATGTGCAATGTGCGAGCTCAGAACATGCACTTGGCTCAGA 3840
Db 3781 CAGGGCCAAAGGCGGGAATGTGCAATGTGCGAGCTCAGAACATGCACTTGGCTCAGA 3840
Qy 3841 CACTGCAAAAGCTCCCTGAGAGCAAGAACTGAAAGAAATGTAATGAACTTCCAGCAT 3900
Db 3841 CACTGCAAAAGCTCCCTGAGAGCAAGAACTGAAAGAAATGTAATGAACTTCCAGCAT 3900
Qy 3901 GGCATCTCCAGTTTCTCTCACTGAGCAACCAATACACAGCGTGAAGGACTTGGGGTGG 3960
Db 3901 GGCATCTCCAGTTTCTCTCACTGAGCAACCAATACACAGCGTGAAGGACTTGGGGTGG 3960
Qy 3961 GTGACAGCTTCTTCCAAACCACTCACTCCCACTGAACGAGAGAGGGGACACTGACTTC 4020
Db 3961 GTGACAGCTTCTTCCAAACCACTCACTCCCACTGAACGAGAGAGGGGACACTGACTTC 4020
Qy 4021 ACCTTCTTCTCGAGAGCTGTTTCACTTCTCAGACCGCGGGCATGCGAGATGCGCATC 4080
Db 4021 ACCTTCTTCTCGAGAGCTGTTTCACTTCTCAGACCGCGGGCATGCGAGATGCGCATC 4080
Qy 4081 GCACTCTGGAACAATGCTGGAACCAAGTGGGCGGCAAGACTTACCTCAACAACTTCAAC 4140
Db 4081 GCACTCTGGAACAATGCTGGAACCAAGTGGGCGGCAAGACTTACCTCAACAACTTCAAC 4140
Qy 4141 CACAGCCGAGCCAAACTCAAGTGCCTCTCTGAGAGCCCTTACCTTACACCTTGCAG 4200
Db 4141 CACAGCCGAGCCAAACTCAAGTGCCTCTCTGAGAGCCCTTACCTTACACCTTGCAG 4200
```

Db	Qy	LOCUS	DEFINITION	SEQUENCE	1 from Patent WO0231161.	DNA	Linear	PAT	18-JUN-2002
Db	Qy	AX417817	Sequence 1 from Patent WO0231161.	AX417817	GI:21522937				
Db	Qy	AX417817	Homo sapiens (human)	Homo sapiens					
Db	Qy	AX417817	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.						
Db	Qy	AX417817	Regulation of human phospholipase-like enzyme						
Db	Qy	AX417817	Patent WO 0231161-A 1 18-APR-2002;						
Db	Qy	AX417817	Bayer AG (DE)						
Db	Qy	AX417817	Location/Qualifiers						
Db	Qy	AX417817	1. 3648						
Db	Qy	AX417817	/organism="Homo sapiens"						
Db	Qy	AX417817	/mol_type="genomic DNA"						
Db	Qy	AX417817	/db_xref="taxon:9606"						
Db	Qy	AX417817	899 a 1032 c 947 g 770 t						
Db	Qy	AX417817	79.1%; Score 3461.2; DB 6; Length 3648;						
Db	Qy	AX417817	Best Local Similarity 97.5%; Pred. No. 0;						
Db	Qy	AX417817	Matches 3558; Conservative 0; Mismatches 33; Indels 57; Gaps 2;						
Db	Qy	AX417817	775 GAAGCTTGAACAGCCTCTGGCCCTCCAGACGTACAGTGAACAGAGCTCTTCAACCGT 834						
Db	Qy	AX417817	1 GAAGCTTGAACAGCCTCTGGCCCTCCAGACGTACAGTGAACAGAGCTCTTCAACCGT 60						
Db	Qy	AX417817	835 GTTTTCCAGCCTTTCTTTATGAGACCAACCCCATCTTCACTCGAAGAACCCCGACTC 894						
Db	Qy	AX417817	61 GTTTTCCAGCCTTTCTTTATGAGACCAACCCCATCTTCACTCGAAGAACCCCGACTC 120						
Db	Qy	AX417817	895 CAGATTTCACACGCTGAGCTGAGCATCTTGGAATAGAGTGAAGACACGACGAGAG 954						
Db	Qy	AX417817	121 CAGATTTCACACGCTGAGCTGAGCATCTTGGAATAGAGTGAAGACACGACGAGAG 180						
Db	Qy	AX417817	955 AAAGTAGCATTGATGATTAACACAGGAGGCCAATGAAGTCCCTTCAGAGAGC 1014						
Db	Qy	AX417817	181 AAAGTAGCATTGATGATTAACACAGGAGGCCAATGAAGTCCCTTCAGAGAGC 240						
Db	Qy	AX417817	1015 CCTTATCTGTTCAAGTACAGAAACAGCACTTACCTGACACAGCTGACAGAAACCCCAAC 1074						
Db	Qy	AX417817	241 CCTTATCTGTTCAAGTACAGAAACAGCACTTACCTGACACAGCTGACAGAAACCCCAAC 300						
Db	Qy	AX417817	1075 AAGCTTG-----AGTAAAGAGA 1092						
Db	Qy	AX417817	301 AAGCTTGAGAAAAATCATTTTCTTATTTTCCCTGTGTCTCATTTCAAGTAAAGAGA 360						
Db	Qy	AX417817	1093 GAGCGGAATACAGTGTCTGACAAAGACCCCTCGATGAGGTTCCACCTCAGTTTCA 1152						
Db	Qy	AX417817	361 GAGCGGAATACAGTGTCTGACAAAGACCCCTCGATGAGGTTCCACCTCAGTTTCA 420						
Db	Qy	AX417817	1153 AGCGTGAAGCGCGGCTGACATCAAGTATTGAGCCCTGGGTGACTCTCTCAACGACAGC 1212						
Db	Qy	AX417817	421 AGCGTGAAGCGCGGCTGACATCAAGTATTGAGCCCTGGGTGACTCTCTCAACGACAGC 480						

OY	1213	AATGGGGCCGGGATCCACACCTGGGGAGCGTCTGGAGCGCTTGAATCACTCAAGTACCGAGGCGCTG	1272
Db	481	AATGGGGCCGGGATCCACACCTGGGGAGCGTCTGGAGCGCTTGAATCACTCAAGTACCGAGGCGCTG	540
OY	1273	TCCTGGAGCGCTGGCGGAGATGAGAACATCGGCACCGTTAACACCTCGTGGCAATCCTC	1332
Db	541	TCCTGGAGCGCTGGCGGAGATGAGAACATCGGCACCGTTAACACCTCGTGGCAATCCTC	600
OY	1333	CGGGAATTCAACCCCTCCCTGGAAAGGCGCTCTCTGTGGCACTGGGAAAGAAACAAGTCC	1392
Db	601	CGGGAATTCAACCCCTCCCTGGAAAGGCGCTCTCTGTGGCACTGGGAAAGAAACAAGTCC	660
OY	1393	AATGCTTCTTTAAACCAAGCTGTGGCAAGGCGCGAGCTGGAGATCTACCTGTCCAGGCC	1452
Db	661	AATGCTTCTTTAAACCAAGCTGTGGCAAGGCGCGAGCTGGAGATCTACCTGTCCAGGCC	720
OY	1453	AGGAGCGCTGTGGACCTGATGGAAGATGACACGAGATACACTTTCAGAAAGACTGGAG	1512
Db	721	AGGAGCGCTGTGGACCTGATGGAAGATGACACGAGATACACTTTCAGAAAGACTGGAG	780
OY	1513	ATAATTAACCCCTGTTATAGCGGCAATGACCTCGTATTTCTGCAATGATCTGTCCAC	1572
Db	781	ATAATTAACCCCTGTTATAGCGGCAATGACCTCGTATTTCTGCAATGATCTGTCCAC	840
OY	1573	TATTTCTCCCCAGAACTTCAACAGCAATTTGAAAGGCCCTTGACATCTTCCATGCTGAG	1632
Db	841	TATTTCTCCCCAGAACTTCAACAGCAATTTGAAAGGCCCTTGACATCTTCCATGCTGAG	900
OY	1633	GTTCTCTGGGGGATTTTGGAAACCTGGTGAACGCGTGTAGATTCGCAACCTGAGGGAGCTG	1692
Db	901	GTTCTCTGGGGGATTTTGGAAACCTGGTGAACGCGTGTAGATTCGCAACCTGAGGGAGCTG	960
OY	1693	TACCAAGGAAAAAGTCTACTGCGCCAAAGATGATTCCTCAAGTCTCTGTGCTCCGTGTC	1752
Db	961	TACCAAGGAAAAAGTCTACTGCGCCAAAGATGATTCCTCAAGTCTCTGTGCTCCGTGTC	1020
OY	1753	CTGAAGTTTGAATGATTACTCAACAGAACTTGCTACCTCATCGAATTTCAACAAAGTTT	1812
Db	1021	CTGAAGTTTGAATGATTACTCAACAGAACTTGCTACCTCATCGAATTTCAACAAAGTTT	1080
OY	1813	CAGAGAAAGACCCACCAACTGATTTGAGAAGTGGCGATATGACACAAAGGAAAGATTTTAACT	1872
Db	1081	CAGAGAAAGACCCACCAACTGATTTGAGAAGTGGCGATATGACACAAAGGAAAGATTTTAACT	1140
OY	1873	GTGTTGTGACGCCGTTCTTTGAAAACGTGACATGCCAAAGACTCTGGAAAGATTGCTT	1932
Db	1141	GTGTTGTGACGCCGTTCTTTGAAAACGTGACATGCCAAAGACTCTGGAAAGATTGCTT	1200
OY	1933	GACAACTCTTCTTGGCTCGTGAATGTTTCCACTTCAAGACAGTGTCACTCCCGAGCA	1992
Db	1201	GACAACTCTTCTTGGCTCGTGAATGTTTCCACTTCAAGACAGTGTCACTCCCGAGCA	1260
OY	1993	GCCAGTGTCTCTTGGAAACAATATGCTGGAGCTGTGGCCAGAAAGACGACTGTCAATAG	2052
Db	1261	GCCAGTGTCTCTTGGAAACAATATGCTGGAGCTGTGGCCAGAAAGACGACTGTCAATAG	1320
OY	2053	TTTGGAAAAACAAGATCAATATTCACATGTGCCAACCAAGGTCCAGCGCTTTCGAGGAACTTAC	2112
Db	1321	TTTGGAAAAACAAGATCAATATTCACATGTGCCAACCAAGGTCTTTCGAGGAACTTAC	1380
OY	2113	AAGAACAAG-----CATGACAGGATCATGGGACCTTGAGCTGCAATGAGGAG	2157
Db	1381	AGATTCTGGAGATTGTCTATGCCCTTATAGGGGTATATGGACCTGTGCTGCAATGAGGAG	1440
OY	2158	AGAGCCCCCTTCTGCTTTGCAACCTTCACTAGTGCATSCCTTGAGACTTGACAGATCTCAA	2217
Db	1441	AGAGCCCCCTTCTGCTTTGCAACCTTCACTAGTGCATSCCTTGAGACTTGACAGATCTCAA	1500
OY	2218	GTTGTGGCTGCTGGGGGGAATTCCTGACGCGTGGCAATGGAAATTTGGCTCCAAACACAGAC	2277
Db	1501	GTTGTGGCTGCTGGGGGGAATTCCTGACGCGTGGCAATGGAAATTTGGCTCCAAACACAGAC	1560

QY 2278 GACCTCCCGGATGTCACACAGATCGGGAGCTGTCAATCAAGTCAGAGAGGGGAGCGC 2337
 DB 1561 GACCTCCCGGATGTCACACAGATCGGGAGCTGTCAATCAAGTCAGAGAGGGGAGCGC 1620
 QY 2338 TCCCTGGAGATGAGACCACTTAATATCTTGGGAGTTTAAACAAAACCTCAC 2397
 DB 1621 TCCCTGGAGATGAGACCACTTAATATCTTGGGAGTTTAAACAAAACCTCAC 1680
 QY 2398 GGCCTACGCGGTGGGACGAGTGTGATGCAATGACAGAAATGATCTTCAATCAAGCTTT 2457
 DB 1681 GGCCTACGCGGTGGGACGAGTGTGATGCAATGACAGAAATGATCTTCAATCAAGCTTT 1740
 QY 2458 CCGGAGCAAAAGGCTGAGATCTTATGAGCAAGTCCAAACTGTATGACAGAAATGAAA 2517
 DB 1741 CCGGAGCAAAAGGCTGAGATCTTATGAGCAAGTCCAAACTGTATGAGAAATGAAA 1800
 QY 2518 GATGATCATGAGTAAATTTCCATGAAAGCTGAAAGTCAACAGTGTGATGAGAGC 2577
 DB 1801 GATGATCATGAGTAAATTTCCATGAAAGCTGAAAGTCAACAGTGTGATGAGAGC 1860
 QY 2578 AGCGATTTATGTAAGTACTGACAGATTCGAATCTGTATCTGAGCAACTTTGTTGAC 2637
 DB 1861 AGCGATTTATGTAAGTACTGACAGATTCGAATCTGTATCTGAGCAACTTTGTTGAC 1920
 QY 2638 CATCTCCGCAATGCTTGAAGTCTGATGATGAGAGGTCGCAAGTCTGTCAACCTC 2697
 DB 1921 CATCTCCGCAATGCTTGAAGTCTGATGATGAGAGGTCGCAAGTCTGTCAACCTC 1980
 QY 2698 GTGGACTTCTGAAACCCCACTATATGCGGAGAGTGTCTTGAGAAAACCAAGTGC 2757
 DB 1981 GTGGACTTCTGAAACCCCACTATATGCGGAGAGTGTCTTGAGAAAACCAAGTGC 2040
 QY 2758 CCGATGAGAGGCGAGCGTTTGTGTAATGAGTCTGACCTTGCGGAGAACTCCCA 2817
 DB 2041 CCGATGAGAGGCGAGCGTTTGTGTAATGAGTCTGACCTTGCGGAGAACTCCCA 2100
 QY 2818 GAGCTAGCGAGCTGAGAGGCTTCAAGCCAGCTTACCGAGAGCAATGCGAGCTGTG 2877
 DB 2101 GAGCTAGCGAGCTGAGAGGCTTCAAGCCAGCTTACCGAGAGCAATGCGAGCTGTG 2160
 QY 2878 GGGTCAGGCGGCTATGACAGCAGAGAGACTTCTGTGTGCTGACGCTTCTTCAG 2937
 DB 2161 GGGTCAGGCGGCTATGACAGCAGAGAGACTTCTGTGTGCTGACGCTTCTTCAG 2220
 QY 2938 AACATCAGCTCCCTCTGCGGAGTGGGCTCCAGATACGCTCTTCTTGGCCAGAC 2997
 DB 2221 AACATCAGCTCCCTCTGCGGAGTGGGCTCCAGATACGCTCTTCTTGGCCAGAC 2280
 QY 2998 TGCATCCACCAATCAGAAATTCACATCCAGCTGGCCAGAGCCCTTGGACCAATATG 3057
 DB 2281 TGCATCCACCAATCAGAAATTCACATCCAGCTGGCCAGAGCCCTTGGACCAATATG 2340
 QY 3058 CTTGAAACCTTGGAAAGCAAGAGACCTTGAAGCTGAGAGAGAGTGGCCATCAC 3117
 DB 2341 CTTGAAACCTTGGAAAGCAAGAGACCTTGAAGCTGAGAGAGTGGCCATCAC 2400
 QY 3118 TGTCCCACTCAGAAATGAGCCCTTCTGAGAAACCCCTTGAATGTAACTACAGTACCC 3177
 DB 2401 TGTCCCACTCAGAAATGAGCCCTTCTGAGAAACCCCTTGAATGTAACTACAGTACCC 2460
 QY 3178 ATCAAGCCAGCCATTTAGAACTGGGAGTGAATCTTCTGTGTACAGAGTGAAGGCTTC 3237
 DB 2461 ATCAAGCCAGCCATTTAGAACTGGGAGTGAATCTTCTGTGTACAGAGTGAAGGCTTC 2520
 QY 3238 AATAGTGTTCACACTTGTGTCACAGCTTCCACAGACAGACATCAAGTGTGAGCGCC 3297
 DB 2521 AATAGTGTTCACACTTGTGTCACAGCTTCCACAGACAGACATCAAGTGTGAGCGCC 2580
 QY 3298 CTGGGAGTCTTGTGACTACAGAGTGGAGCTGACCAAAACATCTCAAGTCACTAAC 3357
 DB 2581 CTGGGAGTCTTGTGACTGTGAGTATGGAGCTGACCAAAACATCTCAAGTCACTAAC 2640
 QY 3358 ACATCTTGAAGGAGCTCTTGTGAGATTTGAGGGAGTGAAGTCTTGAAGTCACTAAC 3417

DB 2641 ACATCTTGAAGGAGACTCTTGGAGCATTTGAGGGGATGGGAACTTGGAGATCTACACC 2700
 QY 3418 ACATCTTGAAGGAGACTCTTGGAGCATTTGAGGGGATGGGAACTTGGAGATCTACACC 3477
 DB 2701 ACATCTTGAAGGAGACTCTTGGAGCATTTGAGGGGATGGGAACTTGGAGATCTACACC 2760
 QY 3478 TGGAGGGAGACAGACAGACTTAAATGTGGCAGCGGAGGGGACAGCTAGGGACATGCCA 3537
 DB 2761 TGGAGGGAGACAGACAGACTTAAATGTGGCAGCGGAGGGGACAGCTAGGGACATGCCA 2820
 QY 3538 GCCCAGGCTGGGACCTGTGAGCGGAATGAAAACAGCCCGACATCAACTGGAGAAA 3597
 DB 2821 GCCCAGGCTGGGACCTGTGAGCGGAATGAAAACAGCCCGACATCAACTGGAGAAA 2880
 QY 3598 GACTGGAAGTGTGACACTCTTCAATTTGGGGTCAAGACTTGTGTCAATTTGTGAGAA 3657
 DB 2881 GACTGGAAGTGTGACACTCTTCAATTTGGGGTCAAGACTTGTGTCAATTTGTGAGAA 2940
 QY 3658 CCGAGAGCCCACTTGGACAGGAATATGTTCAAGCATCCACAGGCGCTGACATCTC 3717
 DB 2941 CCGAGAGCCCACTTGGACAGGAATATGTTCAAGCATCCACAGGCGCTGACATCTC 3000
 QY 3718 TCTGAGAGCTCCCAAGGCTTTCGTCAACGTGTGAGAGTCAATGAGCTGCTAGCTC 3777
 DB 3001 TCTGAGAGCTCCCAAGGCTTTCGTCAACGTGTGAGAGTCAATGAGCTGCTAGCTC 3060
 QY 3778 TACAGGAGCCAAAGCGGGAATGTGCAATGCTGAGCTGACAACTGCACTTGCCTC 3837
 DB 3061 TACAGGAGCCAAAGCGGGAATGTGCAATGCTGAGCTGAGCACTGCACTTGCCTC 3120
 QY 3838 AGACACTGGCAAAAGCTCCCTGAGAGCAAGAACTGAAGAAAGTGAATGGAACTCCAG 3897
 DB 3121 AGACACTGGCAAAAGCTCCCTGAGAGCAAGAACTGAAGAAAGTGAATGGAACTCCAG 3180
 QY 3898 CATGCACTCCAAAGCTTCTCTTCACTGAGCACTGACCAATACACAGGCTGAGACTTGGCGTT 3957
 DB 3181 CATGCACTCCAAAGCTTCTCTTCACTGAGCACTGACCAATACACAGGCTGAGACTTGGCGTT 3240
 QY 3958 GTGTGACAGCTTCTTCCAAACACACTCACCTGAAAGAGAGGGGACACTGAC 4017
 DB 3241 GTGTGACAGCTTCTTCCAAACACACTCACCTGAAAGAGAGGGGACACTGAC 3300
 QY 4018 CTCACTTCTTCTCGAGAGCTGTTTCACTTCTGACCGGGGAGATCCGAGATGGCC 4077
 DB 3301 CTCACTTCTTCTCGAGAGCTGTTTCACTTCTGACCGGGGAGATCCGAGATGGCC 3360
 QY 4078 ATGCACTCTGGAACCAATGCTGAAACAGTGGGCGGCAAGCTCTCAACACTTC 4137
 DB 3361 ATGCACTCTGGAACCAATGCTGAAACAGTGGGCGGCAAGCTCTCTCAACACTTC 3420
 QY 4138 ACCCAGGCGGAGCCAACTCAAGTGCCTTCTCTGAGAGCCCTTACCTTCAACCTG 4197
 DB 3421 ACCCAGGCGGAGCCAACTCAAGTGCCTTCTCTGAGAGCCCTTACCTTCAACCTG 3480
 QY 4198 CGGAACAGCGGATGTCTCCCAACAGAGCTGAAAGAGCCCGAGGTGCTTACTGGGCT 4257
 DB 3481 CGGAACAGCGGATGTCTCCCAACAGAGCTGAAAGAGCCCGAGGTGCTTACTGGGCT 3540
 QY 4258 GTCCAGTGGCAGCGGAGTGGGCTTGTGTGGGATCATGCGGAGCACTGTGGAG 4317
 DB 3541 GTCCAGTGGCAGCGGAGTGGGCTTGTGTGGGATCATGCGGAGCACTGTGGAG 3600
 QY 4318 TGCAGAGAGGTGGCCGAGGGAAGATCTCAATGAGCTGCGCACT 4365
 DB 3601 TGCAGAGAGGTGGCCGAGGGAAGATCTCAATGAGCTGCGCACT 3648

RESULT 3
 OCPHILIP
 LOCUS
 DEFINITION O.cuniculus mRNA for phospholipase.
 ACCESSION Z12841
 4612 bp mRNA linear MAR 09-JUN-1994

VERSION 212841.1 GI:1689
 KEYWORDS phospholipase.
 SOURCE Oryctolagus cuniculus (rabbit)
 ORGANISM Oryctolagus cuniculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 REFERENCE 1
 AUTHORS Boll, W., Schmid-Chanda, T., Semenza, G. and Mantei, N.
 TITLE
 JOURNAL
 MEDLINE
 PUBMED 93286138
 PUBMED 8509424
 2 (bases 1 to 4612)
 AUTHORS Boll, W.
 TITLE
 JOURNAL
 REFERENCE Submitted
 TITLE Direct Submission
 JOURNAL Submitted (25-JUN-1992) WERNER BOLL, Dept. of Biochemistry 11,
 Swiss Federal Institute of Technology, ETH Zentrum,
 Universitaetsstrasse 16, Zurich, CH-8092, Switzerland
 FEATURES
 source
 1..4612
 Location/Qualifiers
 /organism="Oryctolagus cuniculus"
 /mol_type="mRNA"
 /strain="New Zealand White"
 /db_xref="taxon:9986"
 /clone="padrab-1113"
 /cell_type="Epithelial cell"
 /tissue_type="Small intestine"
 /dev_stage="Adult"
 25..4401
 /codon_start=1
 /product="Phospholipase"
 /protein_id="CAA78303.1"
 /db_xref="GI:1690"
 /db_xref="SMISS-PROT:O05017"
 /translation="MALMPSTVFLGLPLLRGADQIOTSSGRNTLRGOLPESLKTFR
 PPEDPKTLASVSESVSHLRPSDIFKVAIIGVERTIPDGAADLEBODTEKRPDQ
 ACWGVVYVLSDIIGFSPSALMPCPELRLVPRGGAEDLMQATELVRSKMEVPOIDPF
 EHDWKLINVFESNTSOCPSPSAQDKGLVIGMDKLRTTLDYILOEVPKAFVNLVDLS
 ELAFLSRWROGAQLSPAEPQRCLETSQTLTKVLTQMSVLAQMSLASKYNTQSESE
 AVAPQPFYESSLSALAEPLDQPTTLALSLMRMMEPIGRKEPSEKEXKRLDCEP
 TOESDYLFTYNSGQLPRVSOQGLAEVREGTEIRCDKPDSDSPVSHRLKPADIKI
 VIGAMGDSLTUNGNGSGQNIIDVLTQYRGLSNGVGDQNIYTTVTLANTLRFPNPS
 LQGSVGTGRTTQAFENQAVAGARAGLLPQAFRLVALMKNDRIINFDQWKIITV
 FIGNDGLCDPCNDPVRSPQNTDNIILHAELPRAVNIYVKLEISKRLVQ
 ETVKSPMIIRLSLCPVLKPEDNSTEIASLETIKAEHQVTOOLIDSGRVDLDDPT
 VILDPFEKVMMPKTDQGLPNSFPAPCPHFSSKHAHAASALMMNMLBPVQKTHH
 NDFEGAVNITCPNQVWPFVLTYSKNSVQSGFTWLCRDRSPASPSPTSHALRPADIV
 VVALGDSLTAGIGIGSKPNDLSGDTQYRGISYSGGDSIDNTVTLNLRQNSNI
 MGFVAGTGDASGTGAFFNQAVGAKARGLMQVQTLVQRMKDHRVNEQDWKIIYVQ
 IGASPLCDQCYTDSNLVYSANFYDHLRLDLALHREVPALVNLVDFPNMVSYTRPOV
 NPDQVQOASALCNCLVSPRENSYEILRLALAAQVSSALRELVESGRVYTRDFSG
 VILDPFEKVMMPKTDQGLPNSFPAPCPHFSSKHAHAASALMMNMLBPVQKTHH
 DLPATVITCTPQNEPFLTRNSDTPPSKPAVYENMGSDPLCTANMLASRVPNSVH
 LQPGIITKVAALGDSLTILMAGARBSNDSPPMFWKGSWSIGDGALETHTTNLNLIK
 KENPSILGFSFTGLEGTGLNVVAVOGARQADMPQAKDLVRMRNSPEIDLEKMKVL
 TLFGNDGLCHFTCEPNEGSSGEYVOHIOQALDPLVYELPTFPVNVSEVMEIADHOD
 OQGRATCLIAAOSHCTCFKYSQSVEMQELKKVMNMLQSGLSRLSYSHQVQYRDPDA
 VQVPEFQNTLVLPLNGRGDTLIPESDDCFHRSERGHAMALANMNLPEVGHCTSN
 NPTYSRTKLKCPSPDSPVLYTLTRNSRLLPDQAEADPTVLYNAVVPAAAGLLIGILMM
 VAGRMRCPREDPEPLSTGL"
 25..81
 sig_peptide
 polyA_signal
 polyA_site
 BASE COUNT 1031 a 1414 c 1293 g 874 t
 ORIGIN
 Query Match 71.4%; Score 3124; DB 4; Length 4612;
 Best Local Similarity 82.6%; Pred. No. 0;
 Matches 3613; Conservative 0; Mismatches 750; Indels 9; Gaps 3;

Db	25	ATGGCTCTGTGGCGGAGCGCTTCTCTCTGGGGGTGTGCGGCTCTGGGGCCAGGTCGA	84
Qy	61	CTTCAGATCCATACTCTCTAGAAAGTACATTGGAAGGCGACTATGGCCAGAGACC	120
Db	85	GATCAGATCCAAAACCTCTTCTGGAAAGAACACATTGGAAAGGCGACTCTGGCCAGAGAGC	144
Qy	121	CTGAAAGATTCTCATTTCCCATGCACACCCAAATATAATTAGAGAGTAATAGCCTCTTAA	180
Db	145	CTGAAGCTTTTCCATTCCATGCGACCCAAAGACGTTACAGAGAGCGTGCCCTCTGAG	204
Qy	181	TCAGTTACTCTCTGAAGCCTTCTGATATTATAAATTTGTGCAGCCATTGGCAATCTGAA	240
Db	205	TCAGTTACTCTCTGAGGCTTCTGACATTAAATTTGTTCAGGCAATTGGCAATGTGAA	264
Qy	241	ATTCTTCAGAACCCAGGGAGCGGCGGATCTGGAAGACAGAACCTGGACTG--AAAGCCA	297
Db	265	ACTGCTCCAGACTCAGGGGCGGAGACCTAGAGAGACAAACGGAGCGGAGAAAGAGCCCA	324
Qy	298	CAGAGAGTGTGCATGGGAGTGAATGACAGTCTCTTACGACATCAATCAATATTTTCAGTCT	357
Db	325	GAGAGGCGTGCATATGGGCGTGGTGAACCGTCTTTCAACATCAATCGGCGCTTCAGTCT	384
Qy	358	TCGTGTTCAATGCGCTGTGTGCCACACTGGAAGAGAGTCAATCCCAAGATGGTGCTGAA	417
Db	385	TCGTGTTGAATGCCGCTGTGTC--CTGAACAAGACTTGTACTCTGTGGAGGTGCTGAA	441
Qy	418	GACTTGTGATTCAGGCTCAAGAACTGTGTAGAAACATGAAAGAGAACTTGCACTTGAC	477
Db	442	GACTTGTGATGAGGCTTACAGAGCTGTGTAGAAAGCATGAGAGAGAACCGCAACTTGAC	501
Qy	478	TTTCAATTGACGTGAAGCTCATCAATGTGTCTTTCAGTAAATGCAGCCAGTGTACTGT	537
Db	502	TTTGAATATACATGGAAGCTCATCAAGTGTCTTTCAGTAAACAAACAGTGTTCCTCG	561
Qy	538	TGCCCTCTGTCTCAACAGAATGGGCTTGCAGGCGGCGCGCTGATGTAGTGGGSGTG	597
Db	562	TGTCTCTCTGTCTCAACGAAGGGGCTGTGTCTGGGGGGCATGGAACAGTGAACAAGAAC	621
Qy	598	CTGCACTACCTGCAGCAGAGAGTCCCGCAGAGCAATTGTAAACCTGTGTGACTTGTGAG	657
Db	622	CTGCACTACCTGCAGCAGAGAGTCCCGAAGCGTTGCTCAACTGTGTGTGATCTCTCGAG	681
Qy	658	GTTGCAGAGGTCTCTGCTCAGTATCAGCGCACTTGGCTCAGCCCTGCAACAGAGCCCTGT	717
Db	682	TTGGCAGACATTTCTCTGCTGCGCTCAGGGGGCTCAGCTCAGCCCAAGCAGCGAACCTTGC	741
Qy	718	AATTGCTCAGAGGAGACACCCGGCTGGCCAAAGTGTGTATGACGTCTTATCAGAAA	777
Db	742	AGTGCTTAAAGGGAACCTTCCAGCTACCAAGTCTTGAACGAGTGTCTTACTTGA	801
Qy	778	GCCTGAACAGCCTCTCTGGGCTTCCAGCAGGTAACGTGAGCAGAGTCTTCAACCGTGT	837
Db	802	GCCTGGGACAGCCTCTCTGGGCTTCCAGGTAACAAACACAGAGGTCTTCTGCGGTGTC	861
Qy	838	TTCCAGCCTTTCTTCTATGAGACCACTTCTTACACTGTGAGAACCCCGCACTCCAG	897
Db	862	TTCCAGCCTTTCTTCTATGAGAGCTCCGTGTGGCGCTGTGGGCGAGCCCGCACTCCAG	921
Qy	898	GATTCTACACGCTGGCTGGCAATCTGTGAAATGTGATGTGAGCGACAGAGAGAGAAA	957
Db	922	GACCCCAACACATGGCGCTGAGTCTTGTGAACGAATGTATGGAGCCCATAGGGCGCAAA	981
Qy	958	GATGAGCAATTGAGTGAATAACAGGAGGCAATGAAGTGTCCCTCTCAGAGAGGCCCC	1017
Db	982	GAAAGCCGTTCACTAGTAGAAGAGAGGAAGCACTGAGATGTCCCAACACAGAGAGGCCCC	1041
Qy	1018	TATCTGTTACGCTACAGAAAACAGCACTACTGACAGCTGCAGAAACCCCAAGACAG	1077
Db	1042	TACCTATTCACTACAGAAACAGGCGCCACTGTCACAGGGTGTTCACAGCCCCAGGGGAAG	1101
Qy	1078	CTTGAAGTAAAGAAAGCGGGAATTCAGATGTCTGTACAAAGACCCCTCCGATACGGTT	1137

Db 1102 CTCGAGGTGAGAGAGGAGCGAAATCCGCTGCCGACAGAGACCCCTCCGACTCGGTT 1161
Qy 1138 CCCACCTCAGTTATAGGCTGAGCGGCTGACATCAACGTAAATGGAGCCCTGGGTGAC 1197
Db 1162 CCCACCTCAGTTATAGGCTGAGCGGCTGACATCAACGTAAATGGAGCTGCAATGGGTGAC 1221
Qy 1198 TCTCTCAGGCGAGCAATGAGGCGGCTGACACCTGGGAAAGCTTTGACGCTTGAAT 1257
Db 1222 TCCCTCAGCGGCAAGGAGCGGCGGCTGCCAGCTGGGAAACATCTCGAATGCTTAACT 1281
Qy 1258 CAGTACCGAGGCTGCTCTGAGCGCTGGGCGGATGAGAACATCGGACCGTTACACC 1317
Db 1282 CAGTACCGAGGCTGCTCTGAGCGCTGGGCGGATGAGAACATCGGACCGTTACACC 1341
Qy 1318 CTGGGCAACCTCTCGGGAATTCACCCCTCCCTGAGAGGCTTCTGTTGGCACTGGG 1377
Db 1342 CTGGGCAACCTCTCGGGAATTCACCCCTCCCTGAGAGGCTTCTCTCGGCACTGGG 1401
Qy 1378 AAGAAACAGCTCTAATGCTTCTTAAACAGGCTGGCAGAGGCGGAGCTGAGAT 1437
Db 1402 AGAGAGCACTTCCAGGCTTCTTCAACAGGCACTGGCGGCGCGGACCGATGGG 1461
Qy 1438 CTACTCTCAGGCGCAGAGGCTGGTGGACCTGATGAGAAATGACAGAGGATACCTT 1497
Db 1462 TTAAATCCCAAGGCGGCTGGTGGCTGATGAGAAATGACAGAGGATAAATTTT 1521
Qy 1498 CAGGAAGCTGGAAGATATACCTGTTATAGGCGGCAATGACCTCTGTGATTTCTGC 1557
Db 1522 CAGGAAGATTTGAAATATATACCGTGTTCATTTGAGGCAACGACCTGTGTATTTCTGC 1581
Qy 1558 AATGATCTGGTTCACATTTCTCCCGAATCTTCAAGACAACTTGGAAAGGCGCTGGAC 1617
Db 1582 AATGATCCGGTGGCTACTCTCCCGAATCTTCAAGACAACTTGGAAAGGCGCTGGAC 1641
Qy 1618 ATCTTCATGCTGAGGTTCTCGGCAATTTGTGAACCTGTGTGACGCTGTGATGCTC 1677
Db 1642 ATCTTCATGCTGAGGTTCTCGGCAATTTGTGAACCTGTGTGAGGCTGTGATGCTC 1701
Qy 1678 AACCTGAGGAGCTGTATCCAGGAGAAAGTCTACTGCCAAGATGATCTCAGGCTT 1737
Db 1702 AACCTGAGGAGCTGTATCCAGGAGAAAGTCTACTGCCAAGATGATGATCTCAGGCTT 1761
Qy 1738 CTGTGTCCTGTCCTGGAATTTGATGATCAACAGACCTGCTACCTCATGAA 1797
Db 1762 CTGTGTCCTGTCCTGGAATTTGATGATCAACAGACCTGCTACCTCATGAA 1821
Qy 1798 TTCGAAGAAGTTTCAAGAGAACCAACCACTGATTGAGGTGGCGATGACACA 1857
Db 1822 ACGATCAAGAGATATCAAGAGAGAGCCAGCAGCTGATTGACAGCGGCGATACAGACG 1881
Qy 1858 AGGGAAGATTTTCTGTGTTGTCAGCGGCTCTTGAAGAGTGAACATGCGAAGACC 1917
Db 1882 AGGGAAGATTTTCTGTGTTGTCAGCGGCTCTTGAAGAGTGAACATGCGAAGACC 1941
Qy 1918 TCGGAAGATTTGCTGCAACCTTTCTTGCTGCTGACGTTTCACTTCACTGACAGAG 1977
Db 1942 CAGGAGGCTTACCCGACAACTTTCTTGCGCGCGACCTGCTTCACTTCACTGACAGAG 2001
Qy 1978 TCTCACTCCGAGCAGCAGTGTCTCTGGAACAAATGCTGAGGCTGTTGGCGAAG 2037
Db 2002 GCCACGCGCAGCGAGTACGCGCTCTGGAACAACTGCTGGAACCTGTTGGCGAAG 2061
Qy 2038 ACGACTGTCTAATGTTGAAAACAGATCAATATCACTGTCGAACAGGCTCAGCGG 2097
Db 2062 ACAACACACATATTTTGAAGAGCAGTCAATTAAGTCCAAACAGGCTGCGCGG 2121
Qy 2098 TTTCTAGAGCCTTCAAGAACAGATGCAAGGCTGAGGAGCTGGCGCTGCAAGAGAG 2157
Db 2122 TTTCTAGAGCCTTCAAGAACAGATGCAAGGCTGAGGAGCTGGCGCTGCAAGAGAG 2181
Qy 2158 AGAGCCCTTCTGCTTGAACCTTACTCAGTGCATGCGCTGAGACCTGAGACATTCAA 2217
Db 2182 AGATCCCGTCTGCTGCGCCCACTCACTGATGACCGCTGAGACCGCGAGATATCAA 2241

Qy 2218 GTTGTGGCTGCTGGGGGATTTCTGACCGCTGGCAATGGAATTGGCTCAACAGAG 2277
Db 2242 GTTGTGGCTGCTGGGAATTTCTTAACTGCTGGCAATGGAATTGGCTCAACAGAG 2301
Qy 2278 GACCTCCCGAGTGCACACAGATCGGGGACTGTACATACAGTCAGAGGGGAGCGG 2337
Db 2302 GACCTCTCAGATGGACCACTCAGTATCGGGGACTGTCAATATAGTTCAAGAGGGAGCGG 2361
Qy 2338 TCCCTGGAAGATGTACCACTTACTTAATTCCTTGGGAGTTTAAACAGAACTTCACA 2397
Db 2362 TCCCTGGAAGATGTACCACTTACTTAATTCCTTGGGAGTTTAAACAGAACTTCATG 2421
Qy 2398 GGCTACCGCTGGGAGCGGGTGTGTCATGACAGAAATGATTCCTCAATCAAGGCTT 2457
Db 2422 GGCTTCCGCTGGGAGCTGTAATGTCAGTGGCAACAAATGCTTTCTTCAACAGGCTTT 2481
Qy 2458 CCCGAGCAAAAGCTGAGGATCTTATGAGCCAAATGCCAACTCTGATGCAAGATGAA 2517
Db 2482 CCCGAGCAAAAGCTGAGGATCTTATGAGCCAAATGCCAACTCTGATGCAAGATGAA 2541
Qy 2518 GATGATCATAGATTAATTTCCATGAAGCTGGAAGTCACTCAGTGTGATGAGAGG 2577
Db 2542 GACGACCATAGATTAATTTCCAAAGATGGAAGTCACTCAGTGTGATGAGAGG 2601
Qy 2578 AGCGATTTATGATCACTGCAAGATGGAATCTGATTTGCAAGCACTTGTGAC 2637
Db 2602 AGCGATTTATGATCACTGCAAGATGGAATCTGATTTGCAAGCACTTGTGAC 2661
Qy 2638 CATCTCCGCAATGCTTGAAGCTCTGCAATGAGAGGTCGCCAGAGTCTGTGCAACTC 2697
Db 2662 CATCTCCGCAATGCTTGAAGCTCTGCAATGAGAGGTCGCCAGAGTCTGTGCAACTC 2721
Qy 2698 GTGGAATTTCTGAACCCCACTATCATGCGGCAAGTGTCTTGGGAAACCCAGCAAGTGC 2757
Db 2722 GTGGAATTTCTGAACCCCACTATCATGCGGCAAGTGTCTTGGGAAACCCAGCAAGTGC 2781
Qy 2758 CCAATGACAGGCGGCTTGTGTATCTGCTTCACTGCTGAGGAGAACTCCCA 2817
Db 2782 CCAATGACAGGCGGCTTGTGTATCTGCTTCACTGCTGAGGAGAACTCCCTAT 2841
Qy 2818 GAGCTAGGCGGCTTGAAGCTTCAAGCGGCTTCAAGAGAGCAATGCGGAGCTGTG 2877
Db 2842 GAGCTAGGCGGCTTGAAGCTTCAAGCGGCTTCAAGAGAGCAATGCGGAGCTGTG 2901
Qy 2878 GAGCTAGGCGGCTTGAAGCTTCAAGCGGAGGACTTCTGTGTGTGTGAGGCTTCTTCA 2937
Db 2902 GAGCTAGGCGGCTTGAAGCTTCAAGCGGAGGACTTCTGTGTGTGTGAGGCTTCTTCA 2961
Qy 2938 AACATCAGCTCTGCTGCGGAGTGGGCTTCCAGATACGCTCTTCTTGGCCGAGAC 2997
Db 2962 AACATCAGCTCTGCTGCGGAGTGGGCTTCCAGATACGCTCTTCTTGGCCGAGAC 3021
Qy 2998 TGCATCAGCAGCAATGAGAAATTCATCTCCAGCTGGCAGAGCCCTTGGACCAATATG 3057
Db 3022 TGCATCAGCAGCAATGAGAAATTCATCTCCAGCTGGCAGAGCCCTTGGCAGCAATATG 3081
Qy 3058 CTTGAACCACTTGAAGCAAAACAGAACCTTGAACCTGAGCCTGAGAGCAGAGATGCCATCAC 3117
Db 3082 CTGGAAGCCCTGGAGAGGAGAAACGAGTGCATGGAACCTGACCGGCAACATCACCTCAC 3141
Qy 3118 TGTCCACTGAGATAGAGCTTCTGAGAACCTCTGGAAATGTAACTACACATACCCC 3177
Db 3142 TGTCCACTGAGATAGAGCTTCTGAGAACCTCTGGAAATGTAACTACACATACCCC 3201
Qy 3178 ATCAAGCCAGCAATGAGAACTGGGGCAGTGAATCTCTGTATCAGAGTGAAGGCTTCC 3237
Db 3202 AGCAGCCAGCCGTGAGAACTGGGGCAGTGAATCTCTGTATCAGAGTGAAGGCTTCC 3261
Qy 3238 AATAGTGTTCACACTTGTTCACAGCTTCGACAGCAGACATCAAAATGTGTGGCGCG 3297
Db 3262 AGGAGTGTTCGGAACCTCACTCATGAGCTCAACAGAGATATCAAAATGTGTGTGGCG 3321

QY 3298 CTGGTGACTCTCTGACTACAGACAGTGGAGCTGACCAACAACTCCAGTGAACCTTACC
 DB 3322 CTGGGTGACTCGGACGCTGGCCATGGGGCTGACCAAGCACTTCCATGACCCACC
 QY 3358 ACATCTTGGAGGGAGCTCTCTTGGAGCATTTGAGGGGATGGGAACTTGGAGACTCACACC
 DB 3382 ATGTTCTGGAGGGGCTCTCTGAGCATTTGAGGGGATGGGAACTTGGAGACTCACACC
 QY 3418 ACATCTGCGCAACATTTGAGAAAGTTCAACCTTACCTTGGCTTCTCTACACAGACC
 DB 3442 ACATCTGCGCAACATTTGAGAAAGTTCAACCTTACCTTGGGGTTTCTTACAGAAC
 QY 3478 TGGAGAGGAGCAGACGAGCACTAAATGTGACGAGGAGGGGCAAGCTAGGAGCATGCCA
 DB 3502 TTGGAGAGGAGCACTGAGCACTCAAGTACGAGTGAAGGGGCAAGCTAGGAGCATGCCA
 QY 3538 GCCGAGGCTGGGAGCTGTGAGAGCAATGAAAACAGCCCGGACATCAACTTGGAGAAA
 DB 3562 GCCCAAGCCCGGAGCTGTGAGAGCAATGAAAACAGCCCGGAGATCGACTTGGAGAAAG
 QY 3598 GACTGGAAGCTGTGACACTTTCATTTGGGGTCAACGACTGTGTCTTACTGTGAGAAAT
 DB 3622 GACTGGAAGCTGTGACACTTTCATTTGGGGGCAATGACCTGTGTCTTGTGAGAAAT
 QY 3658 CCGGAGGCGCACTTGGGAGGAGGAGTACGTACAGCACTTCAAGAGGCTTGGAGCTCTG
 DB 3682 CCGGAGGCGCACTTGGGAGGAGGAGTACGTACAGCACTTCAAGAGGCTTGGAGCTCTG
 QY 3718 TCTGAGAGGCTCCCAAGGCTTTCTGTCAACGCTGTGAGAGTCAATGAGCTGAGCTG
 DB 3742 TACGAGAGGCTTCCGAGGAGCTTTGTCAATGTGTGAGAGTCAATGAGCTGAGCTG
 QY 3778 TACGAGGCGCAAGGCGGAGAAATGTGCA---TGTGTGAGCTGAGCAACACTGTGAGCTG
 DB 3802 CACGAGGAGCAAGGCGGAGAAATGTGCAACTGTGTGAGCTGAGCACTGTGAGCTG
 QY 3835 CTCAGACACTCTGCAAACTCTCTGAGAGAGCAAGAACTGAAAGTGAAGTGAAGCTG
 DB 3862 TTCAAAATCTCTGAGACTCTGAGAGAGTGAAGAACTGAAAGTGAAGTGAAGCTG
 QY 3895 CAGCAGGAGCTTCTCACTTCTCTTCTGAGCACTTCACTGAGCACTGAGCTGAGCTG
 DB 3922 CAGAGGCGGCTCTCAGGCTCTCTCTCACTGAGCACTGAGCTGAGCTGAGCTG
 QY 3955 GTTGTGTGAGAGCTTCTCTTCCAAACAACTGACCTGAGCACTGAGCACTGAGCTG
 DB 3982 GTTGTGTGAGAGCTTCTCTTCCAAACAACTGAGCTGAGCTGAGCTGAGCTGAGCTG
 QY 4015 GACCTCACTTCTCTCTGAGAGCTGTTTCTGAGAGCTGAGCTGAGCTGAGCTGAGCTG
 DB 4042 GACCTCACTTCTCTCTGAGAGCTGTTTCTGAGAGCTGAGCTGAGCTGAGCTGAGCTG
 QY 4075 GGCATGCACTTCTGAGAACTGAGCTGAGCACTGAGCTGAGCTGAGCTGAGCTGAGCTG
 DB 4102 GGCATGCACTTCTGAGAACTGAGCTGAGCACTGAGCTGAGCTGAGCTGAGCTGAGCTG
 QY 4135 TTCAACCAAGGCGGAGCACTGAGCTGAGCACTGAGCTGAGCTGAGCTGAGCTGAGCTG
 DB 4162 TTCAACCAAGGCGGAGCACTGAGCTGAGCACTGAGCTGAGCTGAGCTGAGCTGAGCTG
 QY 4195 CTGCGGAGCAAGGCGGAGCTTCTGAGCACTGAGCTGAGCACTGAGCTGAGCTGAGCTG
 DB 4222 CTGCGGAGCAAGGCGGAGCTTCTGAGCACTGAGCTGAGCACTGAGCTGAGCTGAGCTG
 QY 4255 GCTGTCTCAGTGGAGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG
 DB 4282 GCAATGCGGAGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG
 QY 4315 AGGTGCAAGAGAGTGGCTGAGGAGGAGCTTCAATGAGCTGAGCTGAGCTGAGCTG
 DB 4342 AGAGGCAATGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG

RESULT 4
 AF045454
 LOCUS
 DEFINITION
 Cavia porcellus phospholipase B (PLB) mRNA, complete cds.
 VERSION
 AF045454
 KEYWORDS
 GI:3172336
 SOURCE
 Cavia porcellus (domestic guinea pig)
 ORGANISM
 Cavia porcellus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Hysticognathia; Cavidae; Cavia.
 REFERENCE
 1 (bases 1 to 4608)
 Delagebeaudouf, C., Gassama-Diagne, A., Nauze, M., Ragab, A., Li, R.Y.,
 Capdevielle, J., Ferrara, P., Fauvel, J., and Chap, H.
 Rictopic epididymal expression of guinea pig intestinal
 phospholipase B. Possible role in sperm maturation and activation
 by limited proteolytic digestion
 J. Biol. Chem. 273 (22), 13407-13414 (1998)
 MEDLINE
 PUBMED
 9593672
 2 (bases 1 to 4608)
 Delagebeaudouf, C., Gassama-Diagne, A., Nauze, M., Ragab, A., Li, R.Y.,
 Capdevielle, J., Ferrara, P., Fauvel, J., and Chap, H.
 Direct Submission
 Submitted (30-JAN-1998) INSDRM unfile 326, Hopital Purpan, Toulouse
 31059, France
 FEATURES
 source
 1..4608
 /organism="Cavia porcellus"
 /mol_type="mRNA"
 /strain="Cr1: (BFA) BR"
 /db_xref="taxon:10141"
 /tissue_type="intestine"
 1..4608
 /gene="PLB"
 40..4431
 /note="similar to Adrab-B from adult rabbit intestine"
 /product="phospholipase B"
 /protein_id="AAC40129.1"
 /db_xref="GI:3172337"
 /translation="MGLQGVLLVGLLLGQITQIHTSSGKSLTEGOLMPETKKNP
 FSCSPKGLIMPSESVHTLPADIKLAIAGIMETPPSGAVNLTSRTSEKPMRG
 CMGMVVLSDIISHFNPVSLTPCPMBSAAVGGYELRTQAEELVSLKKNPOLDF
 EAVAMRMQGTMLSRPSEACSGSEKSLDTVMQMSYQETWDSVLASSFNDES
 QODKLIINFPFNSALCYCPSAHENGPELMSNMKLAGLHLYHOSVPAFVNLVDF
 OESPPYFTYRNSVQSRLLKROHREGEITRCEDKPSDSTPSVRLKADIKV
 IGLADSLTAGNGAGSRPQNLIDLVLEYGMSIGADNHSVTLTPNLRFPNPSL
 KGSSTGTGKANSVGAFFNOAVAGARAGDILIPARTLVLMKHTSINFEDEKILTVF
 IGNDICDPCSDPTNPSPEFTDNRQALDIHAIRPAFVNVKVLQVNLRELKXD
 SRVSCRLLRLNRCVLIPDNDSTELSLIDINKYQERTHQLIESGRVIREDTFV
 VLQPFPEKVIDIPKTSGLPDNTSPADCFHFSKTHARASALMKMKPEVQKTON
 NFENSGIDILICPNQAPYSLTYKNGIEGHTWLCRERTSASPTSVHARPADVAV
 AALGDSLTAGSGSKRDLADVIYQYRGISYSGDGLMVTTLPIRLRENSMLT
 GYAVGEGDASNTNAPLNOAVPGAKELMSQKTVYQKKDPRINFPNPSIMQVTLN
 GTNDLCNCTDLDVSSANFNHNLALDIILHREVRALVNLVDPNPSIMQVTLN
 PDCPVOQASILCNQVLSRENSYELARDALTRAQSSMRLEVESGRVDTREDFSV
 LQPFNLNIRLPILEDRPDITSFAPDCINPGQFHSQLSRALVNNLEVSPTSVDLD
 LKADILIPCTOEPPLRTPONSQDYPRKPALENQSDPLCTEMFNSVPSVHKL
 OPADIKVVALGDSLTFVAGARASNSDILMSRGLSWSIGDGLAETHHTLPIILK
 FNPSTIRGFTGTLETRAGNVAVEARADMDPQARDIVRMKASREINLEMDMKLIT
 LPIGSDLCHYCDNPNHSAEYVQIIRQALDILYBELRAPINVIDIIMELAGLHOG
 OGCHTALPLAOSTCSLRFPSSEVPIQELKVTNVLQSDMSLSTQEKYONEDAV
 VVQPFQNTLIPLDKLGSTDPPTFSEDCIHFSERGHAENALIMNNMLPEVGHKTFN
 NPTVNTKLCPSSTESPYLTLQNSLSLTVQTEKASGVAPIVSAAGGLVGLIVG
 ILAVLSMSFRKROKSPESVYNAF"
 BASE COUNT
 1158 a 1329 c 1164 g 957 t
 ORIGIN
 Query Match 67.7%; Score 2962; DB 10; Length 4608;
 Best local Similarity 81.4%; Pred. No. 0;

Matches 3511; Conservative : 0; Mismatches 780; Indels 21; Gaps 6;

QY 1 ATGGGCTGGGCGAGGCAATTTCTCTGAGCTGCTGCTCTTCTGGGCGAAGGACC 60
 Db 40 ATGGGCTGGGCGAGGCGCTTCTCTGAGCTGCTGCTCTTCTGGGCGAAGGACC 96
 QY 61 CCTCAGATCCATCTCTCTAGAAAGATACATTTGAGAGGCGCTATGCGCAGAGACC 120
 Db 97 ACCCAATTCACACATCTCTCTGAGAAAGACACATGAGAGCGCATGATGCGCAGAGACC 156
 QY 121 CTGAAAGATTTCTCATTTCCATGACACCCAAATTAATTAAGAGTGAATATGCTTTAAA 180
 Db 157 AAAAAAGAACTTTCTCATTTCTGTAGCCCAAGAGCTGGAGACTGAACATGCTTTCTGAG 216
 QY 181 TCAGTCACTCTCTGAGACCTTCTGATATTAATTTGTCAGCATTTGCAATCTGGA 240
 Db 217 TCAGTCACTCTCTGAGACCTTCTGATATTAATTTGTCAGCATTTGCAATCTGGA 276
 QY 241 ATTCTTCAGACCCAGGAGCGGGCGATCTGAGAGAGCAAGACTGAGACTG---AAGGCCA 297
 Db 277 ACTCTTCAGACCTCTGCTGAGTCACTGAGACCAAGTGAAGAGACTGAGAAAGAACCA 336
 QY 298 CAGCAGGTGTCATGAGAGTGAATGACAGTCTTTCAAGATCATCATGATATTTCAATCT 357
 Db 337 TGCGGGGGGTGATGAGGAATGATGACAGTCTTTCAAGATCATCATGATATTTCAATCT 396
 QY 358 TCTGTTCCATGCTGTGTGCCACATGGAAGAGATCATCCCGACGATGCTGGA 417
 Db 397 TCTGTTCTGCTGCCACATGCTCTCTCTGAGAGAGCGCTGCTGCGGGTGGTGA 456
 QY 418 GACTTGTGATTCAGGCTCAAGAACTGTGAGAAACATGAAAGAGACTGCAACTGAC 477
 Db 457 GAATTAAGAGCTCAGGCTGAAGAACTGTGAGAGCTTGAAGAAAGACCCGCACTTGAT 516
 QY 478 TTTCAATTTGACTGGAAGCTCATCATATGTGTTCTTCAATATGCAAGCCAGTCTTAC 537
 Db 517 TTTCAACAGATTTGAAACTCATCAATGTGTTCTTGAATGCAAGCCGTGTACTCTG 576
 QY 538 TGCCCTCTGCTCAACAGAAATGGGCTTGCGGGCGGCGGTGATGAGTGAATGGGGTG 597
 Db 577 TGCCCTCTGCTCAACAGAAATGGGCTTGCGGGCGGCGGTGATGAGTGAATGGGGATT 636
 QY 598 CTGAGCTACCTGAGCAGAGAGTCCCGAGAGCAATTTGTAACCTGTGAGACTCTCTGAG 657
 Db 637 CTACACTACCTGAGCAGAGAGTCCCGAGAGCAATTTGTAACCTGTGAGACTCTCTGAG 696
 QY 658 GTTGCAGAGGTCTCTGTGATGATCAAGGCACTTGCTCAGC---CCTGCAGCAGAGCC 714
 Db 697 GTGTGGCCATGCTGATGAGATCAAGGGGCTATGCTCAGAGAGCCCTCACCGGAGGCC 756
 QY 715 TGTATTTGCTCAAGAGAGACACCCGGCTGCGCAAGTGTGATGACAGTGTCTTATCAG 774
 Db 757 TGTGTTGCTCAGAGGAGACCTCAAGCTGATACGTTGTAATGACAGTGTCTTATCAG 816
 QY 775 GAAGCCTGGAAGAGCCTCTGAGCTCAGAGGAGACAGTGAAGAGAGTCTTCAACCGTG 834
 Db 817 GAAGCCTGGAAGAGCCTCTGAGCTCAGAGGAGACAGTGAAGAGAGTCTTCAACCGTG 876
 QY 835 GTTTTCAGGCTTTCTTCTATGAGACCACTCTTACACTGAGAGAGACCCCGCATC 894
 Db 877 GTTTTCAGGCTTTCTTCTATGAGAGTGTCAATCTCTGT-----GGAAGAACCCCGTTC 930
 QY 895 CAGGATTTCTACACGCTGCTGAGATCTCTGAAATGATGATGAGCCAGCAGAGAG 954
 Db 931 CAGGACCCCAACACCTGAGCTGAGCTTGAACCAATGATGAGCCAGAGGAGCAG 990
 QY 955 AAGATGAGCATTGAGTGTAAACACGAGGAGCCATGAGTGTCCCTCTCAGAGAGC 1014
 Db 991 AAGATGAGCATTGAGTGTAAACACGAGGAGCCATGAGTGTCCCTCTCAGAGAGC 1050
 QY 1015 CCCTATCTGTTCACTACAGAAACAGCACTGACAGACTGCAAGAACCCCAAGAC 1074
 Db 1051 CCCTATCTGTTCACTACAGAAACAGCACTGACAGACTGCAAGAACCCCAAGAG 1110

QY 1075 AAGCTTGAGTAAGAGAGGCGAAATCAGATGCTCTGACAAAGACCCCTCCGATACG 1134
 Db 1111 CAGCACAAGAGAGAGAGAGGAGCGAAATCAGATGCTCTGACAAAGACCCCTCCGATAC 1170
 QY 1135 GTTCCACCTCAGTTCATAGGCTGAAGCCGGCTGACATCAACGTAATTTGAGCCCTGAGT 1194
 Db 1171 ACTCCACCTCAGTTCATAGGCTGAAGCCGGCTGACATCAACGTAATTTGAGCCCTGAGT 1230
 QY 1195 GACTCTTCACGAGCAGGCAATGGGGCCGGTCCACACTGAGAAACGCTTGAAGCTCTG 1254
 Db 1231 GACTCTTCACGAGCAGGCAATGGGGCCGGTCCACACTGAGAAACGCTTGAAGCTCTG 1290
 QY 1255 ACTCAGTACGAGGCTGCTCTGAGGCTGGGCGAGATGAGAACATCGGACCGTTAC 1314
 Db 1291 ACTGATCCGAGGCTGCTCTGAGGCTGGGCGAGATGAGAACATCGGACCGTTAC 1350
 QY 1315 ACCCTGCGCAACCTCTCCGGGAATTCACCTTCCCTGAAAGGCTTCTGTGGCACT 1374
 Db 1351 ACCCTGCGCAACCTCTCCGGGAATTCACCTTCCCTGAAAGGCTTCTGTGGCACT 1410
 QY 1375 GGGAAAGAAACAGTCTTAATGCTTTTAAACAGGCTGTGGCAGAGGCCGAGCTGAG 1434
 Db 1411 GGTAAAGCAACAGTGAAGAGCTTCTTAAACAGGCTGTGGCAGAGGCCGAGCTGAG 1470
 QY 1435 GATCTACCTGTCCAGGCGAGAGGCTGTGGAACCTGATGAAGATGACAGAGATCAC 1494
 Db 1471 GATTTGATCCCTCAGGCGAGAGGCTGTGGAACCTGATGAAGATGACAGAGATCAC 1530
 QY 1495 TTTCAAGAACTGGAATTAATTAACCTGTTTAAAGCGGCAATGACTCTGTGATTTTC 1554
 Db 1531 TTTGAGGAATTTGGAATTAATTAACCTGTTTAAAGCGGCAATGACTCTGTGATTTTC 1590
 QY 1555 TGCATGATGTGCTCAATATTTCTCCCGAAGCTTCAAGACCAATTTGAAGGCCCTG 1614
 Db 1591 TGCATGATGTGCTCAATATTTCTCCCGAAGCTTCAAGACCAATTTGAAGGCCCTG 1650
 QY 1615 GACATCTCAGTCTGAGGTTCTCCGGCAATTTGTGAACCTGAGAGGCTGTGAGATC 1674
 Db 1651 GACATCTCAGTCTGAGGTTCTCCGGCAATTTGTGAACCTGAGAGGCTGTGAGATC 1710
 QY 1675 GTCAACTGAGGAGCTGTACAGAGAGAAAGAGTCTACTGCCAGAGATGATCTCAGG 1734
 Db 1711 GTTAACCTGAGGAGCTGTACAGAGAGAAAGAGTCTACTGCCAGAGATGATCTCAGG 1770
 QY 1735 TCTCTGTCTCTGTCTCTGAAATTTGATGATTAATCAACAGAACTTGTACCTCATC 1794
 Db 1771 AACCTGTGTCTGTCTCTGAAATTTGATGATTAATCAACAGAACTTGTACCTCATC 1830
 QY 1795 GAATTCACAAAGATTTCAAGAGAGAAAGCCCACTGATTTGAGAGTGGCCGATATGAC 1854
 Db 1831 GATATTAACAAAGATTTCAAGAGAGAAAGCCCACTGATTTGAGAGTGGCCGATATGAC 1890
 QY 1855 ACAAGGAGATTTTCTGAGTGTGTCAGAGCCGTTCTTGAAGAGTGAAGCATGCCAAG 1914
 Db 1891 ACAAGGAGATTTTCTGAGTGTGTCAGAGCCGTTCTTGAAGAGTGAAGCATGCCAAG 1950
 QY 1915 ACCTCGAAGATTTGCTGCAACCTTTCTGCTCTGATGATTTTCACTTCAAGCAGC 1974
 Db 1951 ACCTCGAAGATTTGCTGCAACCTTTCTGCTCTGATGATTTTCACTTCAAGCAGC 2010
 QY 1975 AAGTCTACCTCCGAGCAGCAGTGTCTTGAACCAATATGCTGAGAGCTGTGGCAG 2034
 Db 2011 AAGTCTACCTCCGAGCAGCAGTGTCTTGAACCAATATGCTGAGAGCTGTGGCAG 2070
 QY 2035 AAGACGATGTCATTAAGTTTGAAGAAACAGATCAATTCATGTCGGAACAGGTCCAG 2094
 Db 2071 AAGACGATGTCATTAAGTTTGAAGAAACAGATGATTCATTTGTCCAAACAGGCTTT 2130
 QY 2095 CCGTTTCTGAGACCTTACAGAAACAGCATGAGGCTCATGGAACCTGTGCTCAGTCCAG 2154
 Db 2131 CCATTACTGATGATCTTACAGAAACAGCATGAGGCTCATGGAACCTGTGCTCAGTCCAG 2190

QY	2155	GACAGAGCCCTTCTGCGCTTGCAACCCATACCTCACTGAGATGCGCCGAGACCTGCAAGATC	2214
Dp	2191	GAGGAGACCCCTTCTTGCTCGCTCCACGTCAGTGGATGCTCGAGACCCGCAAGATGTC	2250
QY	2215	CAAGTTGTGCTGCTCTGGGGGATTTCTCTGACCGCTGGCAATGGAAATTTGGCTCCAAACCA	2274
Dp	2251	CGTCTTGTTGGCGGGGTTGGGGGATTTCTCTGACTCTGGCACTGGAAATTTGGCTCAAAACCA	2310
QY	2275	GACGACCTCCCGATGTCACACACAGTATCGGGGACTGTCTATCAAGTGCAAGAGGGGAC	2334
Dp	2311	GGCAGACCTCGCTGATGTCATCACACAGATCGGGGAGCTGTCTTATGTTCCAGGAGGGGAC	2370
QY	2335	GGCTTCCCTGGAGATGTGACCAACCTTACCTTAATATTCCTGGGGATTAAACAGAAACCTC	2394
Dp	2371	GGCTTCCCTGATGAATGTGACCAACCTTACCTTAATATTCCTGGGGATTAAACAGAAACCTC	2430
QY	2395	ACAGGCTTACCGCGTGGGGGACGGGTGATGCCAATACACAGATGTGATTTCTCAATCAAGCT	2454
Dp	2431	ACTGTTATCGGGTGTGGGCACTGTGTATGCCAGCAACCAATGTGATTTCTCAACAGGCC	2490
QY	2455	GTTCCTCCGAGCAAAAGCTGAGGATCTTATGAGCCAAAGTCCAACTCTGATGCAAGAATG	2514
Dp	2491	GTTCCTCGAGCGAAGGCTGGAAGAACTTATGAGCCAAAGTCAAGTCACTGTGATCAGAAATG	2550
QY	2515	AAAGATGATCATTAAGTAAATTTTCCATGABAACGTGGAAAGTCAACACAGTCTGATGGGA	2574
Dp	2551	AAAGATGATCTTAGAATTAATTTTCCACAAACGTGGAAAGTCAACAGTCTGATGGGA	2610
QY	2575	GGCAGCGATTTATGTGACTACTGCAAGATTCGAATCTGATTTCTGACGCCAACTTTGTT	2634
Dp	2611	ACCAATGATTTGTGCAACCACTGCAAGATTTTGAACCTGTATTTCTTACGCCAACTTTTTC	2670
QY	2635	GACCATCTCGGCAATGCTTGTGAGCGTCTGATATGAGAGGTGCCAAGTCCAGATCTGTGTCAC	2694
Dp	2671	AACCATCTCTCAATGCGCTTGTGAGCAATCTTACATAGAGAGGTTCCAGAGCCTTGTGTCAAC	2730
QY	2695	CTCGTGAACCTTCTGNAACCCCACTATCATATGGGGCAGAGGTTCCTGTGGGAAACCCAGACAAG	2754
Dp	2731	CTCGTGAACCTTCTGNAACCCCACTATCATATGGGGCAGAGGTTCCTGTGGGAAACCCAGACAAG	2790
QY	2755	TGCCAGTGCAGAGGCGCAGCGTTTGTGTAACTGCTTGTGACCCCTGCGGGAGAACTCC	2814
Dp	2791	TGCCCGTGCAGAGGCGCAGCATTTTGTGTCACTGCTGTGACCTGTGAGGAGAACTCC	2850
QY	2815	CAAGAGCTAGCCAGGCTGGAGGCTTTCAGCCGAGCTTACCGGAGCAGCATGCGGAGCTG	2874
Dp	2851	TATGAGCTGGCGAAGATGAGACGCGCTTGACCCGAGCTTACAGAGCAGCATGCTGTAGAGCTG	2910
QY	2875	GTGGGGTATAGGCGCTTATGACACGACGAGAGGACTTCTGTGTGTCTGCAACCTTCTTTC	2934
Dp	2911	GTGGAGTATAGGCGCTTATGACACCGGAGAGGACTTCTGTGTGTCTGCAACCTTCTTTC	2970
QY	2935	CAGAACATCAAGCTCCCTGCTGTGGCGGATGGGCTCCAGATATGCTCTTGTGGCCCCA	2994
Dp	2971	CTCAACATCCGGCTTCCCATCTGTGAGAGTGGGCTCCAGAACCTCTCTTGTGGCCCCG	3030
QY	2995	GACTGCACTCAACCCAAATCAGAAATTCACATCCCACTGGGCGCAGAGCCCTTTGACCAAT	3054
Dp	3031	GACTGCACTCAACCCAGGCGCAAGAAATTCACATCCCACTCTGAGAGCCCTTTGGGTCAAT	3090
QY	3055	ATGCTTTGAACCACTTTGGAAGCAAAAGAGACCTGTGACCTGACACGACAGACATCTCCCTG	3114
Dp	3091	ATGCTTTGAACCTGTGTGGAAGCAAACTGATCTCTGTGACCTGACACGACAGACATCTCCCTG	3150
QY	3115	ACCTGTCCCACTCAATAGAGCCCTTCTGTGGAACCCCTCGGAATATGTAATCAACGCTAC	3174
Dp	3151	CCCTGTCCCACTCAAGAGAGGCGCTTCTGTGGAACCCCTCGGAATAGGACATCAACGCTAC	3210
QY	3175	CCCATCAAGCCAGCATTGAGAACTGGGGCAGTGACTTCTGTGTACAGAGTGAAGGCT	3234
Dp	3211	CCCAACCAAGCCAGCATTGAGAACTGGGGCAGTGACTTCTGTGTACAGAGTGAAGGCT	3270
QY	3235	TCCAATAGTGTTCAACTCTGTCCACAGCTCCGACACAGACATCAATAAAGTGTGCC	3294

Db	3271	TCGATGCGCTCCCAACTCAGTTCAACAGCTCCAAACGACGATCAAAAGTGTGCA	3330
Qy	3295	GCCCTGGGTATCTCTGACTACAGCAGTGGAGCTTGACCAAAACAATCCAGTACCTA	3354
Db	3331	GCCCTGGGTATCTCGCTGACTACAGCAGTGGAGCCGACAAAGCAATCCAGGACCTG	3390
Qy	3355	CCCAATCTTTGGAGGGGACCTCTGTGGAGGATTTGGAGGGATGGGAATTTGGAGATCTGAC	3414
Db	3391	CTCATGTCTGGAGGGGGCTCTCTGGAGCATTTGAGAGAAATGGAGCTTTAGAACCCAC	3450
Qy	3415	ACCACTGCGCCAACTTCTGAAGAAGTTCAACCTTACCTCTTGCTTCTTACGACG	3474
Db	3451	ACCACATGCGCCAACTTCTTAAGAAGTTCAACCTTCCATCTTTGATTTCTACTGGT	3510
Qy	3475	ACCTGGAGGGGACAGCAGAGACTTAATGTGGCAGCGGAGGGGCTCAGAGCTAGGAGCATG	3534
Db	3511	ACCTGGAAAGACGCGCAGAGATTCAGTGGCAGTGGAAAGACCAAGCCTCGGGACATG	3570
Qy	3535	CCAGCCCAAGGCTCTGGGACCTGTGTAGACCGAAATGAATAACGCCCGGACATCAACTGGAG	3594
Db	3571	CCAGCCCAAGGCTCGAGACCTGTGTAGAACGAAATGAATCAACGACCGAGATCAACTGGAG	3630
Qy	3595	AAAGACTGAAGCTGTGTACACTCTTTCATTGGGGTCAACGACTTGTGTCAATTAAGTGTAG	3654
Db	3631	ATGAGACTGGAACTGATCACTCCTCTTCAATTTGGCAGCAATAGCTGTGTCAATTAAGTGTAG	3690
Qy	3655	AATCCGAGGCGCACTTTGGCCACGGAATATGTTCAGCAATCCAAAGGCCCTGGACATC	3714
Db	3691	AATCCGAGGAACCACTCGGCGGAGGATATGTCCAGACATCCGACAGGCCCTGGACATC	3750
Qy	3715	CTCTCTAGAGAGCTCCCAAGGGCTTTGTGTAAGTGTGGG --- GGTATGGAGCTGGCT	3771
Db	3751	CTGTATAGAGAGCTTCCAGAGGGCTTTATCAACCTAGTAGACATCATATGAGACTGGCT	3810
Qy	3772	AGCCTGTATACCAAGGGCCAAAGCGGGGAAATG --- TGCATGTGTGGCAGCTCAGAACATGTGC	3828
Db	3811	GGCTGTGACCAAGGGCCAAAGCGGGCAATTGCACTGCGCTGTGCCAGCAGAGACCTGGC	3870
Qy	3829	ACTTGGCTCAGACACTGCGAAAGCTCCCTGGAGAGAGCAAGACTGAAGAAATGAACTGG	3888
Db	3871	AGTTGGCTCCGACACTTCCCGAGCTCTCTGTGTATCCAAAGGCTGGAAGAAAGTGAACCTGG	3930
Qy	3889	AACTTCCAGATGGGATCTCCAGTTTCTCCTACTGGGACCAATATCACAGCGTGAAGGAC	3948
Db	3931	AACTTCCAGATGACATGTCTCAGGTGTCTTACCAAGAAAGTACAGCGAGCGGAGGAC	3990
Qy	3949	TTTTCGGTTGGTGGCAGACTTTTTCCTCAAAACACACTCCACCTGAAAGGAGAGGG	4008
Db	3991	TTTTCGGTTGGTGGCAGACTTTTTCCTCAAAACAGCTCATCCGCTGGACAAACTCGGG	4050
Qy	4009	GACACTGACCTCACCTTCTTCTCCGAGAGCTGTTTTCATTTTCAGACCGCGGCATGCC	4068
Db	4051	AGCACTGACCCCACTTCTTCTCTGAGAGCTGTCTTCAATTTCTCAGAACTGGGCCAGCG	4110
Qy	4069	GAGATGGCATTCGACATCTCTGGAACAAACATGTGGAAACCATGGGGCCGCAAGACTACCTCC	4128
Db	4111	GAGATGGCATTCGACATCTTGGAACAAACATGTGGAAACCATGGGGCCATTAACAACTTTC	4170
Qy	4129	AACAATTCACCAAGCGGAGCAAACTCAAGTGCCTTCTCTGAGAGCCCTTACCTC	4188
Db	4171	AACAATTCACCTTACCAACCAACCAAACTCAAGTGCCTTCTGACTGAGAGCCCTTACCTC	4230
Qy	4189	TACACCTTCGCGAAGCAGCCATTTGCTCCAGACCAAGGCTGAAGAAGCCCGAGGTGCTC	4248
Db	4231	TACACCTTCGAGAACAGTGTGTCACTCCAGTCCAGACTGAATAAAGCCTTGAGTGTAGCC	4290
Qy	4249	TACTGGGCTGTCCAGATGGCAGCGGAGGTGGGCTTGTGTGTGGGCTATATG	4300
Db	4291	CCGGGATTTGTGTAGCAGCAGCAGAGTGTGGCTCTCTGTGTGGGCTCTATAG	4342

RESULT 5

D63648 LOCUS D63648 4590 bp mRNA linear ROD 01-NOV-2000
 DEFINITION Rattus norvegicus mRNA for phospholipase B, complete cds.
 ACCESSION D63648
 VERSION D63648.1 GI:2696235
 KEYWORDS phospholipase B.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (sites)
 AUTHORS Takemori, H., Zolotarev, F. N., Ting, L., Urban, T., Komatsu, T.,
 Hatano, O., Okamoto, M. and Tojo, H.
 TITLE Identification of functional domains of rat intestinal
 phospholipase B/lipase. Its cDNA cloning, expression, and tissue
 distribution
 JOURNAL J. Biol. Chem. 273 (4), 2222-2231 (1998)
 MEDLINE 98113187
 PUBMED 9442065
 REFERENCE 2 (bases 1 to 4590)
 AUTHORS Tojo, H.
 TITLE Direct Submission
 JOURNAL Submitted (22-JUL-1995) Hiromasa Tojo, Osaka University Medical
 School, Molecular Physiological Chemistry; 2-2 Yamada-oka, Suita,
 Osaka 565, Japan (E-mail: htojo@mr-bio.med.osaka-u.ac.jp,
 Tel: 06-879-3283 (ex. 3283), Fax: 06-879-3288)
 FEATURES
 source location/Qualifiers
 1. 4590
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="SD"
 /db_xref="taxon:10116"
 /tissue_type="small intestine"
 36..4388
 /codon_start=1
 /product="phospholipase B"
 /protein_id="BA23813.1"
 /db_xref="GI:2696236"
 /translation="MESMPGVSLVGLLLLLDQSPSQTIGSGGENTSPQVFRLLK
 NESFPCKPKLESLVLSKSVSLPSDILKLVAILGNLETPAGSGVMNEKQSLSE
 ELONVIGIMTALSDIRHNPSPVLMPTCGPGKGTAGHTIADLIIOAKELVRHLK
 NPDLDEKDKMLITVLFNSOCHLCSDDCKRHLKMKHMLSGVLDYLRBPAKV
 NLVDLSEVLTMAOHOETGSPAPRICKSEETITLKSAMOMSGYOAMDLASAEF
 NGRTRAVVQSGFSEVEPLERPSPODSTTLARITNSMSEVGRKDTLNAEKT
 MKGPSQSLFTYRNSNYQARQLKPTGKQMEGTFCTPDPOSDSLPTVHRLRP
 ADIKVIGAMDLSLTAAGSSGPNVLDVLTQYRGLSWSGDEITETVTLNLRP
 IITVFIGNDLCSGNNLAFSPOTFDNITKALDILHAEPVAFNMVSIETPLR
 ELNEPKVSCPMILRSICPCVNLGENSELAOLVERNKOYASGKILRESYDTR
 DDPTVVLQPMFENVNRPRTLEGIPDSFPAPDCTHRNVTTHASALAKMKMLPYGR
 KTRHONELKPIVPCNQTSPLSTTKNSMLGHGTSMSCERKPSLSPSVTLTPA
 DLDVVALAGSVTAAGNISQSGEDLADVTQYRGLSVSAGDPLFLNVTLLPILRF
 NGMLTYSVGTGVSASAVLNQAVGAKENLMSOVQTLIQCKKDITRVNFDQWV
 ITVMIGASDLDECDNSRYSAANFSDHNLADLHKVEPRALVNLVDMNSIIRQ
 VELKNDKCPVNOCTSVLCNCLTFRGDSHLEALAEFTSYOSMLQVSEGYDIRE
 DRESVIQLPFLNRLPILENGNDPSPAPDCILISOKHTQALRLMANMELPKK
 MDLDELKELALACPTKDKPFLTPRNSNTYPIKPAIEWGSDFLCTEGSPSKVPT
 SVHELPSDLKVVAAAGDPLTATGARPSSTSLDTPWRGLSISIGDGLTETHLP
 NIKTKNPSILGSGTGLTENTAGLVAEBARADMPAQOAKLVKKKSTPTNIQED
 NKILITLLIGNNDCLYCEBPENSTREYVYIQHALDIPEELRPVINVVEMLSG
 LHDGKCAMPLAVOKNGSCIKRSQNLAMOBELKXVNGALSSELSTVHYRMORE
 DRAVTVQPPFRNTFVPLDERGGDLTFESDCHFSYRGAENALILANMMLPVGK
 TTSNFTYNTKTKCPSPENPLIVRNSQILLDKAKENSTLVAVPVAAVGLVVG
 ILQMLMRTVRLVQ"

Query Match 61.4%; Score 2687; DB 10; Length 4590;
 Best Local Similarity 77.1%; Pred. No. 0;
 Matches 3340; Conservative 0; Mismatches 970; Indels 21; Gaps 5;
 QY 28 CTGAGCTGCTGCTTCTGGGGCAAGGACCCCTCAGATCCATACCTCTCTAGAAAG 87
 Db 69 CTGCTGCTGCTACTGCTGAGCAAGGCGCTCCCAATTCACGCGCTTCTGAGAA 128
 QY 88 AGTACATTTGGAAGGAGCATATGCGCAGAGACCTTAAGAAATTCATTCATTCAGAAC 147
 Db 129 AACACATCGGAGGCCGAGCAAGTGTTCGACCTTAAGAAATTTTCATTCCTTGGCAG 188
 QY 148 CCAATTAATTTAGAGTGAATATGCTTCTTAATTCAGTTCACTCTGAGAGCTTGTGAT 207
 Db 189 CCAAGAGTTAGACATGAGTGTGCTTCTTAATGTCAGTTCACTCTGAGACCTTCAGAC 248
 QY 208 ATTAAATTTGTGCGACCATTTGGCAATCTGAAATTCCTCAGAGCCGAGGAGCGG-- 265
 Db 249 ATTAAATGCTGCGACGATCGGCAACTGAAATCTTCAGCGCTGCTGAGCGGTG 308
 QY 266 ---ATCTGAGAGCAAGCAAGTGAAG--GCCACAGAGGTGTGATGAGAGTG 318
 Db 309 GTCAACATGAGAAACCTCAACCTCGAGAGGAACTACAGAAATGTGTCAATGAGAAATC 368
 QY 319 ATGACAGCTCTTACAGATCATCAGATATTTAGTCTTCTGTCATGAGCTGTGTC 378
 Db 369 ATGACAGCTCTTACAGATCATCAGATATTTAGTCTTCTGTCATGAGCTGTGTC 428
 QY 379 CACACTGAAAGAGATCATACCCAC--GATGTGCTGAGAGCTGTGATTCAGGCT 435
 Db 429 TCTCTGGAAGGATCTGAGGCGCACACTATATTGCGAAGACTGTGATTCAGGCT 488
 QY 436 CAAAGCTGTGAAACATGAAGAGAACCTGCAACTTCACTTCAATTTGACTGGAAG 495
 Db 489 AAGAGCTGTGAGGATCTGAAAGCAACCCGAACTGATTTTGAGAGGACTGGA 548
 QY 496 CTGATCAATGTCTTCTAGTAATGCAAGCAGTGTACCTGTGCTGCTGCTCAACAG 555
 Db 549 CTGATCACTGTGCTCTTCACTAACAAGCAGGCTGCTGCTGCTGCTGATCAGCAG 608
 QY 556 AATGCGCTGCGCGCGCGCGCTGATGAGCTGATGGGGTGTGCACTACCTGACAGCAG 615
 Db 609 AAGAGCACTGATGAAGCAACATGAGAGATGCTGACGGGTGCTGATTAACCTGACATCGT 668
 QY 616 GAGCTCCCAAGACATTTGTAACCTGTGAGACCTCTGTGAGTGTGACAGGCTCTGCT 675
 Db 669 GAGTCCCAAGACGCTTGTGAAATTTGTGAGATCTCTGTGAGGTTTGAACATGAGCTGAG 728
 QY 676 GACTATCAAGGCACTTGTGCTACCGCTGACCAAGACCCGTGAATTTGCTAGAGAGACC 735
 Db 729 CAGCATCAAGAGCTGTGCTGACCGCTGACCAAGATTTGCAAAATGCTGAGAGAAATA 788
 QY 736 ACCGCGCTGCGCAAGGTGATGAGCTGATCTTATCAGGAGCTGGAACAGCCTCTG 795
 Db 789 ACGAATTTACAGAGCTGTGATGAGCTGATCTTATCAGGAGCTGAGAGATCTCTG 848
 QY 796 GCTTCAGAGGATGAGTGAAGAGCTCTTACCGGTGTTTCCAGCCTTCTCTAT 855
 Db 849 GCTTCAGAGGATGAGTGAAGAGCTCTTACCGGTGTTTCCAGCCTTCTCTAT 908
 QY 856 GAGACCAACCCATCTTACACTGAGAGACCCCGCACTCCAGATTTCTACAGCTGCGC 915
 Db 909 GAA-----GTAAACTACCTTGTGAGAGGCTCTGCGCCAGGATTTCAACAGCTGCGC 962
 QY 916 TGCACTCTGGAATGATGATGAGAGCCGAGAGAGAAATGAGCATTGAGTGA 975
 Db 963 CTGAGATCTGGAATGATGATGAGAGCCGAGAGAGTGAAGATGAGCACTCAAGAA 1022
 QY 976 AAGCAGGAGGCAATGAAGTGTCTCTCAGAGAGCCCTATCTTGTTCAGCTACAGA 1035
 Db 1023 GCAGAGAGAAACATGAAGATGTCTCTCAGAGAGCCCTATCTATCTTACCTACAGA 1082
 QY 1036 AAGAGCACTACTGACAGCTGAGAGAAACCCCAAGCAAGCTTGAAGTGAAGAGCA 1095
 Db 1083 AATAGCACTACTGAGGCGAGAGAGCTGAAGAACTATGAGAAAGTTTGAAGTGAAGAGCA 1142
 QY 1096 GCGGAATCAGATGTCTTGAAGAGAGCCCTCGATACGCTTCCACTGATTCAGG 1155

1143 ACAAAATTCACCTGTCTGACAAAGACCCCTCGAGCTCATCTCCCAACAAGTTCAAGG 1202
1156 CTGAAGCCGGCTGACATCAACGTAAATTGAGCCCTGGGTGACTCTCAAGGAGGGAAT 1215
1203 CTGAGCGCGGCTGACATCAAGATCATGAGCCATGGGTGACTGTGCTACAGGAGGCAAC 1262
1216 GGGGCGCGGTTCACACCTGGGAAAGCTTTGGACGCTTTGACTGAGTCCGAGGCTGTCC 1275
1263 GGGGCAAGGTTCAAGCCCTGGGAAATGTTCTGGATGTTCTTAACTCAATACAGAGCTGTCC 1322
1276 TGAAGCTCGCGGAGATGAGAACATGCGACCGTTACCACTCTGGCAACATCTTCGG 1335
1323 TGGAGTGTGGCGGAGATGAGAACCATGAGACCTGTGACCACTTGAAGCAATCTTCGG 1382
1336 GAATTCACCTCTCTGAAAGGCTTCTGTTGGCACTGGGAAAGAAACAAGTCTTAAT 1395
1383 GAATTCACCTCTCTGAAAGGCTTCTGTTGGCACTGGGAAAGAAACAAGTCTTCGGA 1442
1396 GCTTCTTAAACAGAGCTGTGACAGAGCGGAGCTGAGATCTACCTGTCCAGGCGCAG 1455
1443 GCATCTTCAACAGAGCGGCTGAGAGAGCCAAATCTGATGGCTTACGTCCAGGCGCAAA 1502
1456 AGGCTGTGAGCTGATGAGAAATGACAGAGATACATTTGAGAAAGCTGAGAGATA 1515
1503 AAGCTGTGAGCTGATGAGAGATGACAGCAATTAACCTTCAGAAAGCTGAGAGATA 1562
1516 ATAAACCTGTTTATAGGCGGCAATGACCTCTGATATTCTGCAATGATCTGCTCACTAT 1575
1563 ATACAGTGTGTTATAGGAGGCAAGACCTCTGAGCTCTGCAATTAACCTGAGCGCTTT 1622
1576 TCTTCCAGAACTTCAAGACAAATTTGAAAGGCGCTGACATCTCTCAAGTCAAGTT 1635
1623 TCTTCCAAACCTTCAAGACAAATTTGAAAGGCGCTGACATCTCTCAAGTCAAGTT 1682
1636 CTTGCGGCAATTTGTGAACCTGTGACGAGTGTGATGATCTGCAACTGAGGAGCTGTAC 1695
1683 CCGCGGCGCTTGTGAACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1742
1696 CAGGAGAAAGATCTACCTGCGCAAGATGATCTCTGAGTCTGATGATGATGATGATGAT 1755
1743 AATGAACTTAAAGTCACTGCGCAAGATGATCTCTGAGGAGCTGTGATGATGATGATGAT 1802
1756 AAGTTGATGATTAATCAAGAACTTGTACCTCTCAATCAATTAATCAAGAACTTGTAC 1815
1803 AACTTGTGTGAATCAAGCAAGAACTTGTCCCAATTTGTGAAAGAAACAGGCAATACAG 1862
1816 GAGAGAGCCCAACAAGTATGAGAGTGGCGGATATGACAAAGGAGAAATTTTACTGTG 1875
1863 GAAGAAACTGGAATACTGATGAGAGTGGCGGATATGACAAAGGAGATGATCTTCAAGCTG 1922
1876 GTTGTGAGGCGCTTCTTGAAGAACTGTGACATGCGCAAGAACTGTGGAAGATGGCTGAC 1935
1923 GTCTTCAGGCGCAATGTTGAAAGATGTGTCAATGCGCAAGAACTGTGAGGCGCTTGCAGAC 1982
1936 AACTCTTCTTCTGCTGCTGATGTTTCACTTCAAGCAAGTGTCACTCCGAGCAGCC 1995
1983 AGCTCTTCTTCTTCTGCTGCTGATGTTTCACTTCAATGATCAAGTCACTCCGCTGAGCC 2042
1996 AGTGTCTGTGAAACAATATGCTGTGAGCTGTGAGGCAAGAAAGCACTGCTATAGTTT 2055
2043 ATGCGCTCTGTGAAGAACATGCTGAGAACTGTGAGGCGGCAAGAAACAATCAAGATTTT 2102
2056 GAAAAACAAGATCAATATCAATGCTGAGAACAGGATCCAGCGCTTCTGAGAGACTTAAG 2115
2103 GAATATCAAGGTCTCTATCATGTGTCCAAACAGACCTTCAAGCTTCTGAGAGACCAAG 2162
2116 AACGATGAGGAGTCAATGAGAACCTGCTGCAATGAGAGCAAGGCGCTTCTGAGCTTG 2175
2163 AACGACAACTTGGAGACATGGAATCTTCAATGCTGAGAGAGAGGCGCTTCTGAGCTTCA 2222
2176 CACCTTAATCTGAGTCAATGCTGAGAACCTGAGACATCAAGTGTGAGTGTGCTGAGG 2235
2223 CCAACAACTCTGAGTCAATGCTGAGAACCTGAGACATCAAGTGTGAGGAGCTGCTGAG 2282

2236 GATTTCTGACCGCTGCAATGGAATTTGGCTCCAAACAGACGACCTCCCGATGTCACC 2295
2283 GACTCTGTACTCTGTGCAATGGAATTCAGTCCCAAGAGATGATCTCGTATGTTACC 2342
2296 ACAAGATATGGAGACTGTGATACAGTGCAGAGAGGAGCGGCTTCTGAGAAATGTGACC 2355
2343 ACAAGATATGAGGACTGTGATACAGTGTGAGAGAGGAGCAAGTGTCTGAGAAATGTGACC 2402
2356 ACCTTACCTTAATTTCTTGGAGGTTTAAACAAGAACTTCAAGGCTTACCGCGGAGGAG 2415
2403 ACCTTGGCAAACTCTCGGAGATTTAATGAAATCTCAAGGCTTCAAGTGTGAGAAC 2462
2416 GGTATGCAATGACAGAAATGATTTCTTCAATCAAGTGTTCGGAGCAAGGCTGAG 2475
2463 GGTAGCTCAACTCTGCAAGCGGCTTCTTAAACAAGCTGTCTCTGGGAGAAAGGCTGAG 2522
2476 GATCTTAATGAGCAAGTCCAACTCTGATGCAAGATGAAAGATGATCATAGATTAAT 2535
2523 AACTTGAAGTCAAGTCCAGACTTGAATTCAGAAATGAAATGACACAGAGTGAAC 2582
2536 TTGCAATGAACTGGAAGTCAATCAAGTGTGATGAGAGGAGGAGGATTAATGATAC 2595
2583 TTTCACCAAGACTGGAAGTCAATCAAGTGTGATGAGAGGAGGAGGAGGAGGAGGAGG 2642
2596 TGCACAGATTCGAATCTGATTTCTGAGGCAACTTGTGATGAGGAGGAGGAGGAGGAGG 2655
2643 TGCAGAGATTCGAATCTGATTTCTGAGGCAACTTGTGATGAGGAGGAGGAGGAGGAGG 2702
2656 GAGCTCTGCAATGAGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2715
2703 GAGCTCTGCAATGAGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2762
2716 ACTATCATGAGGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2775
2763 AGTATCATGAGGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2822
2776 GTTGTGATGATGAGGAGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2835
2823 GTCTGTGCAACTGTGTTCTGAGACCCAGGAGAGATTCCTCAATGAGTGTGAGGAGG 2882
2836 GCTTCAAGCGGAGCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2895
2883 GCTTCAAGCGGAGCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2942
2896 AGCGAGAGGAGCTTCTGT 2955
2943 ACCCGAGAGGATTTCTGT 3002
2956 CTGAGGAGGAGGAGT 3015
3003 CTGAGAGATGAGGAGT 3062
3016 AATTTCACTCCAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3075
3063 AAGTTCAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3122
3076 AAAACAGAGAGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3135
3123 AATATGATCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3182
3136 CCGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3195
3183 CCGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3242
3196 AACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3255
3243 AATTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3302
3256 GTTCACAGGCTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3315
3303 GTTCATGAGCTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3362

QY 3316 ACAGCAGTGGAGCTGCACCAACAACTCCAGTACCTACCACTCTTGAGGGGAGCTC 3375
 DB 3363 ACAGCCACGGAGCTGCACCAAGTGAATGAGTCTAGACACCCCTTGAGGGGAGCTC 3422
 QY 3376 TCTTGAGATTTGAGAGGAGTGGGAATCTTGAGACTCAGACCACTGCCCCAATTCTTG 3435
 DB 3423 TCTTGAGATTTGAGAGGAGTGGGAATCTTGAGACTCAGACCACTGCCCCAATTCTTG 3482
 QY 3436 AAGAAATTCACCCCTTACCTCTTGAGCTCTTACAGACCTGAGAGGGGAGCAGAGGA 3495
 DB 3483 AAGAAATTCACCCCTTACCTCTTGAGCTCTTACAGACCTGAGAGGGGAGCAGAGGA 3542
 QY 3496 CTAAATGTGGCAGAGGGAGGAGCAGAGCTAGAGATGAGACATGACAGCCAGGCTTGGA 3555
 DB 3543 TTAAGTGTGGCAGAGGGAGGAGCAGAGCTAGAGATGAGACATGACAGCCAGGCTTGGA 3602
 QY 3556 GTAGAGCGAATGAAAAACAGCCCCGACATTCACCTGAGAAAAACCTGGAAGCTGTGACA 3615
 DB 3603 GTAGAGCGAATGAAAAACAGCCCCGACATTCACCTGAGAAAAACCTGGAAGCTGTGACA 3662
 QY 3616 CTCTTATTTGGGTCACAGCTTGTGATTAATGAGAAATCCGGAGGCCACTTGAGCC 3675
 DB 3663 CTCTTATTTGGGTCACAGCTTGTGATTAATGAGAAATCCGGAGGCCACTTGAGCC 3722
 QY 3676 ACGGAATATTTGAGACATTCACAGAGCCCTGAGACATCTCTTGAGAGAGCTCCCAAG 3735
 DB 3723 AGGGAATATTTGAGACATTCACAGAGCCCTGAGACATCTCTTGAGAGAGCTCCCAAG 3782
 QY 3736 GCTTTGTCAACGTGTGAGAGTCAATGAGAGCTGAGCTGATCCAGGAGCCAGGCGGG 3795
 DB 3783 GTTTTATCAACCTGTGTGAGAGTCAATGAGAGCTGAGCTGATCCAGGAGCCAGGCGGG 3842
 QY 3796 AATATGCGATGC---TGGAGCTCAGAACTGCACTGCACTGCTCAGACCTGCAAGC 3852
 DB 3843 AATATGCGATGC---TGGAGCTCAGAACTGCACTGCACTGCTCAGAACTGCAAGC 3902
 QY 3853 TCCCTGAGAGCAGAACTGAAAGAGTGAATGGAATCTGCAAGATGAGCATCTCAGAT 3912
 DB 3903 CTGATGCGAATGAGAGAGTGAAGAAAGTCAACGGAACTCCAGAGTGCCTCTTGAGAG 3962
 QY 3913 TTCTCTTACTGAGCACAATACACAGAGCTGAGAGCTTTGCGGTTGTGTCAGCCCTTTC 3972
 DB 3963 CTCTCTTACTGAGCACAATACACAGAGCTGAGAGCTTTGCGGTTGTGTCAGCCCTTTC 4022
 QY 3973 TTCCAAACACACTCACCACCTGAGAGAGAGGGGAGCAGTCACTCTCTCTCTC 4032
 DB 4023 TTCCGAAATACCTTTGTCCACTGATGAGCGTGGGGGCTTGCACTCTCTCTCTCT 4082
 QY 4033 GAGGACTGTTTCACTTCTCAGACCGCGGAGCATGCGAGATGAGCATTCGCACTCTGAAAC 4092
 DB 4083 GAGGACTGTTTCACTTCTCAGACCGCGGAGCATGCGAGATGAGCATTCGCGCTCTGAAAC 4142
 QY 4093 AACATGCTGAGAACAGTGGGCGGCAAGACTTACTCTCAACAACTTCAACCAAGCGAGCC 4152
 DB 4143 AACATGCTGAGAACAGTGGGCGGCAAGAAACAACTTCAACAACTTCAACCAAGCGAGCC 4202
 QY 4153 AAATCTAAGTGGCCCTCTCTCTGAGAGCCCTTACTTCAACCTCTGAGAAACAGCCGATTTG 4212
 DB 4203 AAATCTAAGTGGCCCTCTCTCTGAGAGCCCTTACTTCAACCTCTGAGAAACAGCTCGAAT 4262
 QY 4213 CTCCGAGACAGGCTGAAGAAAGCCCGAGAGTCTCTAAGTGGGCTGCTCCAGTGGAGCG 4272
 DB 4263 CTCTTGAACAGGCTGAAGAAAGCTCCAAATACCTTACTGAGGAGTGGCAGTGGCTGCA 4322
 QY 4273 GAGTGGGCTTGTGTGGGAGCATTCAGGAGCAGTGTCTGAGAGTGTGAGAGAGTGGC 4332
 DB 4323 GTAGTGGCCTGTGTGTGGGAGCATTCCTGGAATGATGTGTGTGAGAGCTGTGAGACTGCTC 4382
 QY 4333 CGAGGAGGAGA 4343
 DB 4383 CAATAGAGGA 4393

RESULT 6
 E13935
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 E13935
 Rat mRNA for phospholipase.
 E13935.1 GI:3252702
 JP 1997248190-A/1.
 Rattus sp.
 Rattus sp.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 4613)
 Tojo, H. and Hasegawa, A.
 NEW PHOSPHOLIPASE AND DNA CODING THE SAME
 Patent: JP 1997248190-A 1 22-SEP-1997;
 TOJO HIROMASA, TONEN CORP
 OS
 Rattus sp. (rat)
 PN
 JP 1997248190-A/1
 PD
 22-SEP-1997
 PF
 15-MAR-1996 JP 1996086022
 PI
 TOJO HIROMASA, HASEGAWA AKIRA
 PC
 C12N15/09, C07H21/04, C07K14/47, C12N5/10, C12N9/16, PC
 C12N9/20//A61K38/46,
 CC
 C12N5/10, C12R1.91), (C12N9/16, C12R1.91), (C12N9/20, C12R1.91);
 CC
 strandedness: Double;
 CC
 topology: Linear;
 FH
 Key
 Location/Qualifiers
 FH
 source
 1. .4613
 /organism="Rattus sp."
 FT
 CDS
 36. .4357
 /lign="tissue type='ileum'
 FT
 sig_peptide
 36. .125
 /product="phospholipase precursor" FT
 FT
 mat_peptide
 126. .4394
 /product="phospholipase".
 FT
 Location/Qualifiers
 1. .4613
 /organism="Rattus sp."
 /mol_type="genomic DNA"
 /db_xref="taxon:10118"
 BASE COUNT
 1218 a 1243 c 1142 g 1010 t
 ORIGIN
 Query Match 61.4%; Score 2687; DB 6; Length 4613;
 Best Local Similarity 77.1%; Pred. No. 0;
 Matches 3340; Conservative 0; Mismatches 970; Indels 21; Gaps 5;
 QY 28 CTGAGCTGCTGCTGCTTGTGGGGCAAGGAGCCCTCAGATCCATGCTCTCTAGAAAG 87
 DB 69 CTGCTGCTGCTGCTGCTGCTGAGCAAGGAGCCCTCCAAATCCAGCGCTCTTGTGAGAA 128
 QY 88 AGTACATTGAAGGGAGCTATGAGCCAGAGACCCCTGAGAAATTTCTCATTTCCATCCAAAC 147
 DB 129 AACACATGCGACGCCAGCAAGTGTTCGGAACCCCTGAAGATTTTCATTTCCCTTCAAG 188
 QY 148 CCAATTAATTAAGAGTGAATATGCTTTAATCAGTTCACTCTGTGAAGCTTGTGAT 207
 DB 189 CCAAGAAATTAAGTGAATATGCTTTAATCAGTTCACTCTGTGAAGCTTGTGAT 248
 QY 208 ATTAATTTGTGAGAGCCATTTGGCAATCTGGAATTTCTCCAGACCCAGAGAGGCGC-- 255
 DB 249 ATTAATTTGTGAGAGCCATTTGGCAATCTGGAATTTCTCCAGACCCAGAGAGGCGC 308
 QY 266 ----ATCTGAGAGCAAGACTGAGCTGAAG---GCCACAGAGGTGTGATGAGAGT 318
 DB 309 GTCAACATGAGAGAACTTCAAGCCCTGAGAGCACTACAGAAATGTGTGATGAGAAATC 368
 QY 319 ATGACAGTCTTTTCAAGATCATCAGATATTTCAAGTCTTTGTGTTCAATGCCCTGTGTC 378
 DB 369 ATGACAGTCTTTTCAAGATCATCAGATATTTCAAGTCTTTGTGTTCAATGCCCTGTGTC 428

Db 2583 TTTCACCAAACTGGAAAGTCACTGTATGATTTGGGGCCAGCGACTTGTGTGACTTC 2642
 QY 2596 TGCAGAGATTGCAATCTGTATTCTGACGCCAATTGTGACCATCTCCGCAATGCCCTTG 2655
 Db 2643 TCGAAGATTTCGAATGCTTACTCTGACGCCAATTTTCTGACCATCTCCGCAATGCCCTTG 2702
 QY 2656 GACGTCCTGCAATGAGAGTGGCCAGAGTCTGTGCAACCTGTGTGAATCTTCTGAAACCC 2715
 Db 2703 GACATCTGCTATGAGAGGTACCCAGAGCCCTGGTCACTGTGTGAATCTTCTGAAACCC 2762
 QY 2716 ACTATCATGGGGCAAGTGTCTGTGGAAATCCAGACAGTGGCCAGTGTGACGAGCCAGC 2775
 Db 2763 AGTATCATTCGGCAAGTGTCTGTGAAGAACCCAGACAGTGGCCCTGTGAATCAGACAGT 2822
 QY 2776 GTTTGTGTAACTGCGTTCTGACCTGCGGGAGAACTCCAGAGCTAGCCAGGCTGTGAG 2835
 Db 2823 GTCTCTGCACTGTGTCTGTACCCAGGGAGAGATTCCTCATGAGCTGGGCAAGGTGTGAG 2882
 QY 2836 GCTTTAGCCGAGCTTACCCGAGCAGCATGCGGAGCTGTGTGGGTGTCAGGCGCTATGAC 2895
 Db 2883 GCTTTACCAAAATCTTACAGAGTAGATGCTTCACTGGTTGAGTCAAGGCGCTAGAC 2942
 QY 2896 ACCGAGAGGACTTCTGT 2955
 Db 2943 ACCCGGAGGATTTCTGT 3002
 QY 2956 CTGGCGGATGGGCTCCAGATAGTCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3015
 Db 3003 CTGAGAGATGGGAATCCAGATATCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3062
 QY 3016 AAATTCACCTCCAGCTGAGCCAGAGCCCTTGTGACCAATATGTGTGAACTGTGAAAGC 3075
 Db 3063 AAGTTCAACACTAGCTCGGAGAGCCCTTGTGGCCAAATATGTGTGAAACCCCTGGAAAG 3122
 QY 3076 AAAACAGAGACCTGTGACTGTGAGACAGATGCCATCACTGTGTGTGTGTGTGTGTGTGT 3135
 Db 3123 AAATGTGATCTTGTGAGCCGAAAGAACTCATAGCTTGTGTGTGTGTGTGTGTGTGTGTGT 3182
 QY 3136 CCCTTCGTGAGACCCCTCGGAATAGTAATTAACGTAACCCCATCAAGCCAGCATTTGAG 3195
 Db 3183 CCCTTCGTGAGACCTTCGGAACAGTATCAACGTAACCCCATCAAGCCAGCATTTGAG 3242
 QY 3196 AACTGGGGAGTACTCTGTGTGAGAGTGTGAAAGCTTCAATATGTGTGTGTGTGTGTGTGT 3255
 Db 3243 AATTGGGGAGTACTCTGTGTGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3302
 QY 3256 GTTCACACAGTCTCGACAGCAGACATCAATGATGTGTGTGTGTGTGTGTGTGTGTGTGT 3315
 Db 3303 GTTCATGAGCTCGACATCATGATCATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3362
 QY 3316 ACAGCAGTGGAGTGTGACCAAACTCACTGATGACTTACCCCATTTTGTGAGGGAGCTC 3375
 Db 3363 ACAGCAGTGGAGTGTGACCAAACTCACTGATGACTTACCCCATTTTGTGAGGGAGCTC 3422
 QY 3376 TCTTGGAGCTTGTGAGGGAGTGTGAACTGTGAGACTCAACCACTGAGCCCATTTG 3435
 Db 3423 TCTTGGAGCTTGTGAGGGAGTGTGAACTGTGAGACTCAACCACTGAGCCCATTTG 3482
 QY 3436 AAGAGTTCACCTTACTCTGT 3495
 Db 3483 AAGAGTTCACCTTACTCTGT 3542
 QY 3496 CTAAATGTGCGAGGAGGGGCGAGAGTGTGAGACTGTGAGCTGAGCCAGCTGTGAGCTG 3555
 Db 3543 TTAAATGTGCGAGGAGGGGCGAGAGTGTGAGACTGTGAGCTGAGCCAGCTGTGAGCTG 3602
 QY 3556 GTAGAGGAAATGAAAAACAGCCCGACATCACTGTGAGAAAGCTGTGAGCTGTGAGCA 3615
 Db 3603 GTAGAGGAAATGAAAAACAGCCCGACATCACTGTGAGAAAGCTGTGAGCTGTGAGCA 3662
 QY 3616 CTCTTCATTTGGGCTCAACGACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3675
 Db 3663 CTCTTCATTTGGGCTCAACGACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3722

QY 3676 ACCGATATGTTTCAAGACATCCAAAGCCCTGAGCATCTCTGTGAGAGCTCCCAAG 3735
 Db 3723 AGGAGATATGTCAAGATCATCCAGCATGCTTTGAGCATCTTCTATAGAGAGCTTCCAGG 3782
 QY 3736 GCTTTGTCAACGTGTGTGAGGTATGAGCTGTGAGCTGTGATCAAGGCGCAAGGCGG 3795
 Db 3783 GTTTTCATCAACGTGTGTGAGGTATGAGCTGTGAGCTGTGATCAAGGCGCGG 3842
 QY 3796 AAATGTGCTATGCT--TGGAGCTCAGAAACATGTGACTTGTGCTCAGACACTGTGAAAGC 3852
 Db 3843 AAATGTGCTATGCT--TGGAGCTCAGAAACATGTGACTTGTGCTCAGACACTGTGAAAGC 3902
 QY 3853 TCCCTGAGAGCAAGAACTGAAAGAAAGTGAATGTGAACTCCAGATGAGCATCTCCAGT 3912
 Db 3903 CTATGAGCAATGAGAGCTGTGAAAGAAAGTCAACGGAACCTCCAGATGAGCATCTCCAG 3962
 QY 3913 TTCTCTACTGTGAGCAATACACAGCGTGAAGACTTGTGCGGTGTGTGTGTGTGTGTGTGT 3972
 Db 3963 CTCTCTACTGTGAGCAATACACAGCGTGAAGACTTGTGCGGTGTGTGTGTGTGTGTGTGT 4022
 QY 3973 TTCCAAACACACTACCCCATCTGAACGAGAGGGGACACTGACTCAGCTTCTTCTCC 4032
 Db 4023 TTCCGAAATACCTTGT 4082
 QY 4033 GAGAGCTGTTTCACTTCTGAGACGCGGAGCATGCGGAGATGAGCATGAGCATCTGAGAAC 4092
 Db 4083 GAGAGCTGTTTCACTTCTGAGACGCGGAGCATGCGGAGATGAGCATGAGCATCTGAGAAC 4142
 QY 4093 AACATCTGGAACAGTGTGGGCGGAGACTTCTTCAACAACTTCAACCCAGCCAGGCC 4152
 Db 4143 AACATCTGGAACAGTGTGGGCGGAGACTTCTTCAACAACTTCAACCCAGCCAGGCC 4202
 QY 4153 AAATCAAGTGTGCTTCTGT 4212
 Db 4203 AAATCAAGTGTGCTTCTGT 4262
 QY 4213 CTCCGACACAGCTGAAGAGCCCGGAGGTGCTGTACTGTGGGCTGTCCAGTGTGAGCG 4272
 Db 4263 CTTCAGACAGGCTTAAGAAATCTCAATATCACTTACTGTGGGAGGCGAGTGTGCA 4322
 QY 4273 GGAGTGGGCTTGT 4332
 Db 4323 GTAGTGTGCTGT 4382
 QY 4333 CGAGGGGAAAG 4343
 Db 4383 CAATAGAGGA 4393

RESULT 7
 AX711962 1835 bp DNA linear PAT 10-Apr-2003
 LOCUS AX711962
 DEFINITION Sequence 1 from Patent WO02062977.
 ACCESSION AX711962
 VERSION AX711962.1 GI:29787748
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 Yan, C., Ketchum, K., di Francesco, V., and Beasley, E.M.
 Human phospholipase b-like polypeptide and uses thereof
 Patent: WO 02062977-A 1 15-AUG-2002;
 PE Corporation (NY) (US)
 LOCATION/Qualifiers
 1..1835
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

FEATURES
 source
 458 a 549 c 474 g 354 c

BASE COUNT
 ORIGIN

Query Match 31.1%; Score 1362; DB 6; Length 1835;
 Best Local Similarity 89.3%; Pred. No. 0;
 Matches 1582; Conservative 0; Mismatches 0; Indels 190; Gaps 3;

QY 2618 CTGACGCACTTTGTTGACATCTCGCAATGCTTGAAGCTTCATAGAG----- 2671
 DB 1 CTGACGCACTTTGTTGACATCTCGCAATGCTTGAAGCTTCATAGAGCTTT 60
 QY 2672 -----AGTGGCCAGATCTGTGTCACCTCGTGAATCTTGAACCCCATATGCTGC 2725
 DB 61 TCCCTTAGGAGCCAGAGCTGTGTCACCTCGTGAATCTTGAACCCCATATGCTGC 120
 QY 2726 GGGAGGTGTTCTGAGAAACCCAGACAAGTGCAGAGCCAGGCTTTGTGTA 2785
 DB 121 GGGAGGTGTTCTGAGAAACCCAGACAAGTGCAGAGCCAGGCTTTGTGTA 168
 QY 2786 ACTGCGTTGACCTTGCGGAGAACTCCAGAGCTAGCCAGGCTTGAGCC 2845
 DB 169 ----- 168
 QY 2846 GAGCGTACCGGAGAGATCGGAGCTGTGGGTGAGGCGGTATGACACGAGAGG 2905
 DB 169 -----GAGCAGATCGGAGCTGTGGGTGAGGCGGTATGACACGAGAGG 218
 QY 2906 ACTTCTGTGTGCTGAGAGCCCTTCTTCAGAACTCCAGCTCCGTGCGGAGATG 2965
 DB 219 ACTTCTGTGTGCTGAGAGCCCTTCTTCAGAACTCCAGCTCCGTGCTGAGC----- 273
 QY 2966 GGTCTCCAGATAGCTCTTCTTGGCCCGAGCTGATCCACCAATCAGAAATTCACCT 3025
 DB 274 ----- 273
 QY 3026 CCCAGCTGGCAGAGCCCTTGGACCAATATGCTTGAACCACTTGAAGCAAAACAGAGA 3085
 DB 274 -----GCTTGAACCACTTGAAGCAAAACAGAGA 302
 QY 3086 CCTGAGCTGAGAGAGATGCGCATCTGCTCCCACTGAGAAATGAGCCCTTCTGTA 3145
 DB 303 CCTGAGCTGAGAGAGATGCGCATCTGCTCCCACTGAGAAATGAGCCCTTCTGTA 362
 QY 3146 GAAACCCCTGGAATAGTAATCTACAGTACCCCATCAAGCCATTTAGAACTGGGCA 3205
 DB 363 GAAACCCCTGGAATAGTAATCTACAGTACCCCATCAAGCCATTTAGAACTGGGCA 422
 QY 3206 GTGACTTCTGTGTGACAGATGGAAGGCTTCCATATGTTCCAACTCTGTCCACAGC 3265
 DB 423 GTGACTTCTGTGTGACAGATGGAAGGCTTCCATATGTTCCAACTCTGTCCACAGC 482
 QY 3266 TCCGACGAGAGACATCAAAAGTGTGCGCCCTGGGTGACTCTGTGACTTACAGCAGTGG 3325
 DB 483 TCCGACGAGAGACATCAAAAGTGTGCGCCCTGGGTGACTCTGTGACTTACAGCAGTGG 542
 QY 3326 GAGCTGACCAAACTCCAGTGAACCTCACTTGAAGGGAATCTCTTGGAGCA 3385
 DB 543 GAGCTGACCAAACTCCAGTGAACCTCACTTGAAGGGAATCTCTTGGAGCA 602
 QY 3386 TTGGAGGGGATGGGAATTTGGAAGTCAACACACTGCCCCACATTTCTGAAGAATTCA 3445
 DB 603 TTGGAGGGGATGGGAATTTGGAAGTCAACACACTGCCCCACATTTCTGAAGAATTCA 662
 QY 3446 ACCCTTACCTCTGTTGCTTCTTACAGCACTGGGAGGGGAGACGAGAGCTAAATGTTG 3505
 DB 663 ACCCTTACCTCTGTTGCTTCTTACAGCACTGGGAGGGGAGACGAGAGCTAAATGTTG 722
 QY 3506 CAGCGAGAGGGGAGAGTGAAGGACATGCGAGCCAGGCTTGGAATCTGTGAGAGGAA 3565
 DB 723 CAGCGAGAGGGGAGAGTGAAGGACATGCGAGCCAGGCTTGGAATCTGTGAGAGGAA 782
 QY 3566 TGAAGAAACAGCCCGACATCACTTGAAGAAAGTGAAGCTGTGACACTTCTTCAATTG 3625
 DB 783 TGAAGAAACAGCCCGACATCACTTGAAGAAAGTGAAGCTGTGACACTTCTTCAATTG 842

QY 3626 GGGTCAACGACTTGTGTACTTGTAGATTCGAGAGGCCCACTTGGCCACGGAATATG 3685
 DB 843 GGGTCAACGACTTGTGTACTTGTAGATTCGAGAGGCCCACTTGGCCACGGAATATG 902
 QY 3686 TTGAGCAGATCCAAAGGCCCTGGAACATCTCTGTGAGAGTCCCAAGGGCTTTGTCA 3745
 DB 903 TTGAGCAGATCCAAAGGCCCTGGAACATCTCTGTGAGAGTCCCAAGGGCTTTGTCA 962
 QY 3746 ACGTGGGAGGATGAGAGCTGCTAGCTGATACAGAGGCCCAAGGGGAGAAATGTGCCA 3805
 DB 963 ACGTGGGAGGATGAGAGCTGCTAGCTGATACAGAGGCCCAAGGGGAGAAATGTGCCA 1022
 QY 3806 TGTGAGAGTCAAGAACTGCACTTGTGAGACCTTGCAAGCTCCCTGAGAGAGC 3865
 DB 1023 TGTGAGAGTCAAGAACTGCACTTGTGAGACCTTGCAAGAGTCCCTGAGAGAGC 1082
 QY 3866 AAGAACTGAAGAAAGTGAATGGAACCTTCAGATAGGATCTTCAGTTTCTTCACTGGC 3925
 DB 1083 AAGAACTGAAGAAAGTGAATGGAACCTTCAGATAGGATCTTCAGTTTCTTCACTGGC 1142
 QY 3926 ACCAATACACAGCGGTGAGAGCTTGGCGGTGAGAGCTTCTTCAGAAACACAC 3985
 DB 1143 ACCAATACACAGCGGTGAGAGCTTGGCGGTGAGAGCTTCTTCAGAAACACAC 1202
 QY 3986 TCACCCCACTGAACGAGAGAGGAGACCTGACCTTCTTCTCGAGAGCTGTTTC 4045
 DB 1203 TCACCCCACTGAACGAGAGAGGAGACCTGACCTTCTTCTCGAGAGCTGTTTC 1262
 QY 4046 ACTTCTGAGACGCGGAGATGCGAGATGGCCATGCACTTGTGAACACATGCTGGAC 4105
 DB 1263 ACTTCTGAGACGCGGAGATGCGAGATGGCCATGCACTTGTGAACACATGCTGGAC 1322
 QY 4106 CAGTGGGCGGCAAGCTACCTCCAAACACTTCCACCAAGCCGAGCAAACTCAAGTGC 4165
 DB 1323 CAGTGGGCGGCAAGCTACCTCCAAACACTTCCACCAAGCCGAGCAAACTCAAGTGC 1382
 QY 4166 CCTCTCTGAGAGCCCTTACCTTACAGCCTGCGGAACAGCCGATTTCTCCAGACAG 4225
 DB 1383 CCTCTCTGAGAGCCCTTACCTTACAGCCTGCGGAACAGCCGATTTCTCCAGACAG 1442
 QY 4226 CTGAAGAGCCCGGAGGTGCTTACTGAGCTTCCAGTGGGAGAGGTGGCTTGG 4285
 DB 1443 CTGAAGAGCCCGGAGGTGCTTACTGAGCTTCCAGTGGGAGAGGTGGCTTGG 1502
 QY 4286 TGTGGGAGATCAAGGGAGAGGTGCTGAGAGTGCAGAGAGGTGCGGAGGGAAAGATC 4345
 DB 1503 TGTGGGAGATCAAGGGAGAGGTGCTGAGAGTGCAGAGAGGTGCGGAGGGAAAGATC 1562
 QY 4346 CTCGAATGAGCTGCGCACTGTGGCCCTCTAG 4377
 DB 1563 CTCGAATGAGCTGCGCACTGTGGCCCTCTAG 1594

RESULT 8
 BC042674 1790 bp mRNA linear PRI 09-JAN-2003
 LOCUS Homo sapiens, similar to phospholipase B, clone MGC:35447
 DEFINITION IMAGE:5191712, mRNA, complete cds.
 ACCESSION BC042674
 VERSION BC042674.1 GI:27503748
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (02-JAN-2003) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>

COMMENT

Contact: MGC help desk
 Email: cgsabs-x@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegod, H.,
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nannavalli,
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAK plate: 50 Row: k Column: 10.

FEATURES

source

CDS

Location/Qualifiers
 1. 1790
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:35447 IMAGE:5191712"
 /tissue_type="Brain, adult, 6 pooled whole brains"
 /clone_lib="NIH_MGC_114"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 196..1467
 /codon_start=1
 /product="Similar to phospholipase B"
 /protein_id="AAH42674.1"
 /db_xref="GI:27503749"
 /translation="MPITCTONEPPLRPNSNYTPRIKPALENKWSDFCTEMKAS
 NSVPSVHOLRPADIKVVAALGSLITAVARPNNSDLPTRSGKSMSTGSGNET
 HTLPIILKFNPIYLAGFSTSTWEGTALNVAAGAPARWPAQAWDLERKNSPDI
 NLEKDWLVTLFGVNDLCHYCNENPAHLATEYVQHIQQALDILSELPSAFVNVVY
 MELASLYOGSGKCAIAAONCTCLRHSGSLSEKQBLKVNMLDHGJLSPFSYVQY
 TOEDPAVVVQPFONTLPFLNERGDTLTFPSEDCFHFEDESRHAEMALALNNMIEP
 VGRKTSNNFTHSRAKLKCPSPSPYLYTLRNSRLLPDQAEAPVLYNAVPVAAGVG
 LVVIGITGVVRRRGRREDPPMSLRVAL"

BASE COUNT

519 a 508 c 437 g 326 t

Query Match 30.1%; Score 1319.4; DB 9; Length 1790;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1320; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3057 GCTTGAACCACTTGAAGCAAAAAGAGACCTGAGACCTGAGAGAGAGATGCCATCAC 3116
 DB 147 GCTTGAACCACTTGAAGCAAAAAGAGACCTGAGACCTGAGAGAGATGCCATCAC 206
 QY 3117 CTGTCCCACTCAGAAATGAGACCCCTTCTGAGAACCCCTCGGAATATGTAACACGTAACC 3176
 DB 207 CTGTCCCACTCAGAAATGAGACCCCTTCTGAGAACCCCTCGGAATATGTAACACGTAACC 266
 QY 3177 CATCAAGCCAGCAATGAGAACTGGGAGAGTCTTCTGTATCAGAGTGAAGGCTTC 3236
 DB 267 CATCAAGCCAGCAATGAGAACTGGGAGAGTCTTCTGTATCAGAGTGAAGGCTTC 326
 QY 3237 CAATAGTGTTCACAACTCTGTCCACAGCTCCGACAGCAGACATCAAAATGTTGCGCG 3296
 DB 327 CAATAGTGTTCACAACTCTGTCCACAGCTCCGACAGCAGACATCAAAATGTTGCGCG 386
 QY 3297 CCTGGGTGACTCTCTGACTAGACAGAGTGGAGCTCGACAAACAATCTCCAGTACCTACC 3356
 DB 387 CCTGGGTGACTCTCTGACTAGACAGAGTGGAGCTCGACAAACAATCTCCAGTACCTACC 446
 QY 3357 CACATCTGGAGGGAGCTCTCTTGGAGCAATTGAGAGGGAGAGGAACCTTGGAGACTACAC 3416
 DB 447 CACATCTGGAGGGAGCTCTCTTGGAGCAATTGAGAGGGAGAGGAACCTTGGAGACTACAC 506
 QY 3417 CACATCTGGAGGGAGCTCTCTTGGAGCAATTGAGAGGGAGAGGAACCTTGGAGACTACAC 3476

DB 507 CACATCTGGAGGGAGCTCTCTTGGAGCAATTGAGAGGGAGAGGAACCTTGGAGACTACAC 566
 QY 3477 CTGGAGAGGGAGAGCAGAGACTAATATGAGAGGGAGAGGGAGAGGAGCTAGGAGCAATGCC 3536
 DB 567 CTGGAGAGGGAGAGCAGAGACTAATATGAGAGGGAGAGGGAGAGGAGCTAGGAGCAATGCC 626
 QY 3537 AGCCAGAGGCTGGAGACTGTATAGAGCAATGAAAAAGAGCCCGAGACATCAACTGGAGAA 3596
 DB 627 AGCCAGAGGCTGGAGACTGTATAGAGCAATGAAAAAGAGCCCGAGACATCAACTGGAGAA 666
 QY 3597 AGACTGAAGCTGTGTACACTCTTCAATGGGGTCAACAGACTGTGTCAATTAAGTGAGAA 3656
 DB 687 AGACTGAAGCTGTGTGTACACTCTTCAATGGGGTCAACAGACTGTGTCAATTAAGTGAGAA 746
 QY 3657 TCCGAGAGGCTCAATTTGGAGAGGAGAAATATGTTGAGCAATCCAGAGCCCTGGAGCAATCCT 3716
 DB 747 TCCGAGAGGCTCAATTTGGAGAGGAGAAATATGTTGAGCAATCCAGAGCCCTGGAGCAATCCT 806
 QY 3717 CTCTGAGAGCTCCCAAGGGCTTTTCTCAACGTGTGTGAGAGTCAATGAGCTGAGCTAGCCT 3776
 DB 807 CTCTGAGAGCTCCCAAGGGCTTTTCTCAACGTGTGTGAGAGTCAATGAGCTGAGCTAGCCT 866
 QY 3777 GTACCAAGGCAAGGCGGGAATGTGCAATGCTGCGAGCTCAGACAACTGCACTTGCTCT 3836
 DB 867 GTACCAAGGCAAGGCGGGAATGTGCAATGCTGCGAGCTCAGACAACTGCACTTGCTCT 926
 QY 3837 CAGACACTGCGAAGAGCTCCCTGAGAGAGCAAGAACTGAAAGAAAGTAACTGGAACCTTCCA 3896
 DB 927 CAGACACTGCGAAGAGCTCCCTGAGAGAGCAAGAAAGTAACTGGAACCTTCCA 966
 QY 3897 GCATGAGCATCTCCAGTTTCTCTCACTGAGCAACATACACAGCGTGAAGACTTTGCGGT 3956
 DB 987 GCATGAGCATCTCCAGTTTCTCTCACTGAGCAACATACACAGCGTGAAGACTTTGCGGT 1046
 QY 3957 TGTGTGAGAGCTTTCTTCCAAAAACACTCACCCCACTGAAAGAGAGGGAGCACTGA 4016
 DB 1047 TGTGTGAGAGCTTTCTTCCAAAAACACTCACCCCACTGAAAGAGAGGGAGCACTGA 1106
 QY 4017 CCTGACCTTCTTCTCCGAGAGCTTTTCACTTCTCAGACCGGCGGCAATGCCAGATGCC 4076
 DB 1107 CCTGACCTTCTTCTCCGAGAGCTTTTCACTTCTCAGACCGGCGGCAATGCCAGATGCC 1166
 QY 4077 CATGCACTCTGGAACAACATGCTGGAACAGAGGAGCCGCAAGACTCACTCCAAACAATT 4136
 DB 1167 CATGCACTCTGGAACAACATGCTGGAACAGAGGAGCCGCAAGACTCACTCCAAACAATT 1226
 QY 4137 CACCCAAGCCGAGCCAAACTCAAGTCCCTTCTCTGAGAGCCCTTAACCTTACACCTT 4196
 DB 1227 CACCCAAGCCGAGCCAAACTCAAGTCCCTTCTCTGAGAGCCCTTACCTTACACCTT 1286
 QY 4197 GCGGAAACAGCCGATTTCTCCAGACAGAGCTGAAGAAAGCCCGAGAGTGTCTTACTGGGC 4256
 DB 1287 GCGGAAACAGCCGATTTCTCCAGACAGAGCTGAAGAAAGCCCGAGAGTGTCTTACTGGGC 1346
 QY 4257 TGTGCCAGTGGAGCGGAGAGTGGAGCTTGTAGTGGAGATCATGAGGAGAGTGTGGAG 4316
 DB 1347 TGTGCCAGTGGAGCGGAGAGTGGAGCTTGTAGTGGAGATCATGAGGAGAGTGTGGAG 1406
 QY 4317 GTGCAGAGAGGTGGCGGAGAGGAAGATCTCCAAATGAGAGCTTGGAGCTTGGAGCTTGA 4376
 DB 1407 GTGCAGAGAGGTGGCGGAGAGGAAGATCTCCAAATGAGAGCTTGGAGCTTGGAGCTTGA 1466
 QY 4377 G 4377
 DB 1467 G 1467

RESULT 9
 AX713706 2477 bp DNA linear PAT 15-APR-2003
 LOCUS AX713706
 DEFINITION Sequence 390 from Patent Epi293569.
 AX713706
 ACCESSION AX713706.1 GI:29888564
 VERSION

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
1 Isogai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahari, K. and Masuno, Y.

TITLE
JOURNAL
Full-length cDNAs
Patent: EP 1293569-A 390 19-MAR-2003;
Helix Research Institute (Jp) ; Research Association for
Biotechnology (Jp)
Location/Qualifiers
1. 2477
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 691 a 597 c 591 g 598 t

ORIGIN

Query Match 22.28; Score 970; DB 6; Length 2477;
Best Local Similarity 99.5%; Pred. No. 5,7e-244;
Matches 973; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1279 AGCGTGGCGGAGATGGAACATCGGACCGCTTACACCTGGGGAACATCTCCGGGAA 1338
|||||
6 AGCGTCGGCGAGATGGAACATCGGACCGCTTACACCTGGGGAACATCTCCGGGAA 65
|||||

1339 TTCAACCCCTTCCCTGGAAGGCTTCTCTGTTGGCACTGGGAAAGAACCGCTTAATGCC 1398
|||||
66 TTCAACCCCTTCCCTGGAAGGCTTCTCTGTTGGCACTGGGAAAGAACCGCTTAATGCC 125
|||||

1399 TTCTTAAACCAAGCTGTGGCAGAGGCGGACGTGAGGATCTACTGTCCAGGCCAGAGG 1458
|||||
126 TTCTTAAACCAAGCTGTGGCAGAGGCGGACGTGAGGATCTACTGTCCAGGCCAGAGG 185
|||||

1459 CTGTGTGACCTGATGGAAGATGACAGAGATACATTTTCAGGAAAGCTGGAAATATA 1518
|||||
186 CTGTGTGACCTGATGGAAGATGACAGAGATACATTTTCAGGAAAGCTGGAAATATA 245
|||||

1519 ACCCTGTTTAAAGCGGCGCAATGACCTGTGATTTTTCGCAATGATCTGTCTCACTATTCT 1578
|||||
246 ACCCTGTTTAAAGCGGCGCAATGACCTGTGATTTTTCGCAATGATCTGTCTCACTATTCT 305
|||||

1579 CCCGAGAACTTTCACAGACCAATTTGAAAGGCGCTGACATCTCCATGCTGAGGTTCT 1638
|||||
306 CCCGAGAACTTTCACAGACCAATTTGAAAGGCGCTGACATCTCCATGCTGAGGTTCT 365
|||||

1639 CGGGCAATTTGTGAACCTGTGACCGTGTCTGAGATCTCTCAACCTGAGAGGCTGTACAG 1698
|||||
366 CGGGCAATTTGTGAACCTGTGACCGTGTCTGAGATCTCTCAACCTGAGAGGCTGTACAG 425
|||||

1699 GAGAAAAAAGTACTACTGCCCAAGGATGATCTCTGAGTCTCTGTTCCCTGTGCTTGAAG 1758
|||||
426 GAGAAAAAAGTACTACTGCCCAAGGATGATCTCTGAGTCTCTGTTCCCTGTGCTTGAAG 485
|||||

1759 TTGTGATGATTAACCAAGAACTTGTACCTCTCATTCGAATTTCAAGAAAGTTTCAGAG 1818
|||||
486 TTGTGATGATTAACCAAGAACTTGTACCTCTCATTCGAATTTCAAGAAAGTTTCAGAG 545
|||||

1819 AAGAACCAACCACTGATTTGAGAGTGGCGCATATGACACAGGAAAGTTTAACTGTGTGT 1878
|||||
546 AAGAACCAACCACTGATTTGAGAGTGGCGCATATGACACAGGAAAGTTTAACTGTGTGT 605
|||||

1879 GTGAGCGCGTTCTTTGAAAAGCTGAGCATGCAAAAGACTCGGAAAGATTGCCGGAAC 1938
|||||
606 GTGAGCGCGTTCTTTGAAAAGCTGAGCATGCAAAAGACTCGGAAAGATTGCCGGAAC 665
|||||

1939 TCTTTCTTCTGCTCTGACTGTTTTCACCTTACAGCAGCAAGTCTCACTCCGAGCAGCCAGT 1998
|||||
666 TCTTTCTTCTGCTCTGACTGTTTTCACCTTACAGCAGCAAGTCTCACTCCGAGCAGCCAGT 725
|||||

1999 GCTCTGGAACAATATGCTGAGACCGTGTGGCCAGAAAGCACTGCTATAAGTTTGA 2058
|||||
726 GCTCTGGAACAATATGCTGAGACCGTGTGGCCAGAAAGCACTGCTATAAGTTTGA 785
|||||

2059 AACAGATCAATATCAATATGCTGGAACCAAGTCCAGCCGTTTCTGAGAGCTTCAAGAAC 2118
|||||
786 AACAGATCAATATCAATATGCTGGAACCAAGTCCAGCCGTTTCTGAGAGCTTCAAGAAC 845
|||||

2119 AGCATGAGAGGTATGAGACCTGCTGCTGCTATGAGGACAGACCCCTTGTGCTTGCAC 2178
|||||
846 AGCATGAGAGGTATGAGACCTGCTGCTGCTATGAGGACAGACCCCTTGTGCTTGCAC 905
|||||

2179 CCTACCTGAGTGCATGCGCTGAGACCTGAGACATCCAAATTTGTGCTGAGGAGAT 2238
|||||
906 CCTACCTGAGTGCATGCGCTGAGACCTGAGACATCCAAATTTGTGCTGAGGAGAT 965
|||||

2239 TCTCTGACCGCTGCAAT 2256
|||||
966 TCTCTGACCGCTGCAAT 983
|||||

RESULT 10
AK055428 2477 bp mRNA linear PRI 01-AUG-2002
LOCUS Homo sapiens cDNA FLJ10866 fis, clone F088A2004110, highly similar
DEFINITION
ACCESSION AK055428.1 GI:16550150
VERSION AK055428.1 GI:16550150
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
1 Nishi, T., Nakagawa, S., Senoh, A., Mizuguchi, H., Inagaki, H., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J. I., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuno, Y., Nagai, K. and Isogai, T.

TITLE
JOURNAL
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2477)

REFERENCE
AUTHORS
1 Isogai, T., Otsuki, T. and Sugiyama, T.

TITLE
JOURNAL
Direct Submission
Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

COMMENT
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'-3' end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

FEATURES
source
1. 2477
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="F088A2004110"
/tissue_type="brain"
/clone_lib="F088A2"
/dev stage="fetus"
/note="cloning vector: pME18SFL3"
198..1010
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB70920.1"
/db_xref="GI:16550151"

/translation="MKNDTRIHEQEDWKIITLFIGNDLDFCNLDVHYSPONFDTNI
 GVALDILHAEPRAFNLYVLEIVNLRELYOEKRYCPRMIDRSCPLVKEDNST
 ELATLEFNKKFOEKTHOLIESGRYDTRBEFTVVOPEFNMIDRSCPLKEDNST
 POCFHSKSHSRAASALNMNMLEPVQKTRHKFNKINITCPOVOPFLIRYKNSM
 OGHGTLPCRDRAPSALHPSVHALRPADIQVALGDLSTLTTLGPVWVG"
 BASE COUNT 691 a 597 c 591 g 598 t
 ORIGIN

Query Match 22.2%; Score 970; DB 9; Length 2477;
 Best Local Similarity 99.5%; Pred. No. 5,7e-244;
 Matches 973; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1279 AGCGTCGGCGAGATGAGAAATCGGACCGTTACACCCCTGCGAATCTCCGAGAA 1338
 DB 6 AGCGTCGGCGAGATGAGAAATCGGACCGTTACACCCCTGCGAATCTCCGAGAA 65
 QY 1339 TTCAACCTTCTCTGAAAGGCTTCTCTGTGGCACTGGGAAAGAACAGCTCTAATGCC 1398
 DB 66 TTCAACCTTCTCTGAAAGGCTTCTCTGTGGCACTGGGAAAGAACAGCTCTAATGCC 125
 QY 1399 TTCTTAACGAGCTGTGGAGAGGCGGAGTGAAGATCTACCTGTCCAGGCGAGAGG 1458
 DB 126 TTCTTAACGAGCTGTGGAGAGGCGGAGTGAAGATCTACCTGTCCAGGCGAGAGG 185
 QY 1459 CTGGTGACCTGATGAAGATGACAGAGGATACATTTGAGGAGAGCTGAAGATAA 1518
 DB 186 CTGGTGACCTGATGAAGATGACAGAGGATACATTTGAGGAGAGCTGAAGATAA 245
 QY 1519 ACCCTGTTATAGGCGGCAATGACTCTGTGATTTCTGCAATGATCTGTCTCAATTC 1578
 DB 246 ACCCTGTTATAGGCGGCAATGACTCTGTGATTTCTGCAATGATCTGTCTCAATTC 305
 QY 1579 CCCCAACCTTCACAGAAATTTGAAAGGCGCTGAGCACTCTCCATGCTGAGTTCCT 1638
 DB 306 CCCCAACCTTCACAGAAATTTGAAAGGCGCTGAGCACTCTCCATGCTGAGTTCCT 365
 QY 1639 CGGGCAATTTGTGAACCTGTGTGACGCTGTGTGATGCTCAACTGAGGAGCTGTACAG 1698
 DB 366 CGGGCAATTTGTGAACCTGTGTGACGCTGTGTGATGCTCAACTGAGGAGCTGTACAG 425
 QY 1699 GAGAAAAAGTCTACCTGCCCCAGAGATGATCTCAGGCTCTGTGTCTCTGTCTGAA 1758
 DB 426 GAGAAAAAGTCTACCTGCCCCAGAGATGATCTCAGGCTCTGTGTCTCTGTCTGAA 485
 QY 1759 TTTGATGATTAACCAAGAACTTGTCTACCTCATGGAATTCAGAAAGTTTCAGAG 1818
 DB 486 TTTGATGATTAACCAAGAACTTGTCTACCTCATGGAATTCAGAAAGTTTCAGAG 545
 QY 1819 AAGACCCCAACCTGATGAGAGTGGGCGATATGACAGAGGAGATTTTACTGTGGTT 1878
 DB 546 AAGACCCCAACCTGATGAGAGTGGGCGATATGACAGAGGAGATTTTACTGTGGTT 605
 QY 1879 GTGACAGCGTCTTTGAAAACGTGACATGCAAGACCTCGAGAGATTTGCTGCAAC 1938
 DB 606 GTGACAGCGTCTTTGAAAACGTGACATGCAAGACCTCGAGAGATTTGCTGCAAC 665
 QY 1939 TCTTTTCTGCTCTGATCTGTTCCTCACTTCACTGAGCAAGTCTCACTCCGAGAGCAG 1998
 DB 666 TCTTTTCTGCTCTGATCTGTTCCTCACTTCACTGAGCAAGTCTCACTCCGAGAGCAG 725
 QY 1999 GCTCTGGAACAATGCTGAGCGCTGTGTGGCAGAGAGCACTGTCTAATGTTGAA 2058
 DB 726 GCTCTGGAACAATGCTGAGCGCTGTGTGGCAGAGAGCACTGTCTAATGTTGAA 785
 QY 2059 AACCAAGATCAATATCATATGTCGAAACAGAGTCAGCGCTTTCTGAGACCTTCAAGAA 2118
 DB 786 AACCAAGATCAATATCATATGTCGAAACAGAGTCAGCGCTTTCTGAGACCTTCAAGAA 845
 QY 2119 AGCATGAGGAGTATGAGAGCTGTGCTGCACTGAGGAGAGAGCCCTTCTGCTTGCAC 2178
 DB 846 AGCATGAGGAGTATGAGAGCTGTGCTGCACTGAGGAGAGAGCCCTTCTGCTTGCAC 905
 QY 2179 CTAACCTCACTGATGAGCCCTGAGACCTGAGACATCAAGTTGTGCTGTGCGGAGAT 2238

|||||
 DB 906 CTAACCTCACTGATGAGCCCTGAGACCTGAGACATCAAGTTGTGCTGCTGCGGAGAT 965
 QY 2239 TCTCTGACCGCTGGCAAT 2256
 DB 966 TCTCTGACCGTAAGACT 983

RESULT 11
 AX417821
 LOCUS AX417821 608 bp DNA linear PAT 18-JUN-2002
 DEFINITION Sequence 5 from Patent WO0231161.
 ACCESSION AX417821
 VERSION AX417821.1 GI:21522939
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE
 1 Zhu Z.
 Regulation of human phospholipase-Like enzyme
 Patent: WO 0231161-A 5 18-APR-2002;
 BAYER AG (DE)
 FEATURES
 source location/Qualifiers
 1..608
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 140 a 187 c 158 g 123 t
 ORIGIN

Query Match 13.2%; Score 576.2; DB 6; Length 608;
 Best Local Similarity 97.8%; Pred. No. 3.2e-140;
 Matches 584; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
 QY 2654 TGAACGTCCTGATAGAGAGTGGCCAGAGTCTGTGCAACTGCTGAGACTTCTGAACC 2713
 DB 11 TGAATGGCTGATAGAGAGTGGCCAGAGTCTGTGCAACTGCTGAGACTTCTGAACC 70
 QY 2714 CCACTATCATGCGGAGGCTTCTTGGGAAACCCAGAACCAAGTGCAGAGGCGCA 2773
 DB 71 CCACTATCATGCGGAGGCTTCTTGGGAAACCCAGAACCAAGTGCAGAGGCGCA 130
 QY 2774 GCGTTTGTGTATCTGCTTCTGACCTTCCGAGAACTCCCAAGCTTACAGGCTG 2833
 DB 131 GCGTTTGTGTATCTGCTTCTGACCTTCCGAGAACTCCCAAGCTTACAGGCTG 190
 QY 2834 AGGCTTCAGCGAGGCTTCCGAGAGATGCGAGGCTGGTGGGCTGAGGCGGCTATG 2893
 DB 191 AGGCTTCAGCGAGGCTTCCGAGAGATGCGAGGCTGGTGGGCTGAGGCGGCTATG 250
 QY 2894 ACAACGAGAGAGACTTCTGTGTGTGCTGACGCTTCTTCCAGAACTACAGCTCC 2953
 DB 251 ACAACGAGAGAGACTTCTGTGTGTGCTGACGCTTCTTCCAGAACTACAGCTCC 310
 QY 2954 TCTGTGCGAGTGGGCTCCAGATACGTCCTTCTTCCAGACTGATCCAAATC 3013
 DB 311 TCTGTGCGAGTGGGCTCCAGATACGTCCTTCTTCCAGACTGATCCAAATC 370
 QY 3014 AGAAATTCACCTCCAGCTGCGAGGCTTGTGAGCAATGTGTTGAACAACCTTGGAA 3073
 DB 371 AGAAATTCACCTCCAGCTGCGAGGCTTGTGAGCAATGTGTTGAACAACCTTGGAA 430
 QY 3074 GCAAAACAGAGACCTTGAACCTGAGAGAGAGATGCCATCTGTCCACTCAGAAATG 3133
 DB 431 GCAAAACAGAGACCTTGAACCTGAGAGAGAGATGCCATCTGTCCACTCAGAAATG 490
 QY 3134 AGCCCTTCTGAGAAACCTTGGAAATGTAATCACTACAGTCCCATCAAGCCAGCTT 3193
 DB 491 AGCCCTTCTGAGAAACCTTGGAAATGTAATCACTACAGTCCCATCAAGCCAGCTT 550
 QY 3194 AGAAGTGGGAGATCACTTCTGTGTACAGAGTGGAGGCTTCAATATGTTCCA 3250

```

Db      551 AGAAGTGGGAGTACTTCTGTGTACAGAAAGTGAAGCTTCAATAGTGTCCAA 607
RESULT 12
LOCUS   AX417820/c
DEFINITION Sequence 4 from Patent WO0231161.
ACCESSION AX417820
VERSION  AX417820.1 GI:21522938
KEYWORDS
SOURCE  Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Zhu, Z.
TITLE    Regulation of human phospholipase-11ike enzyme
JOURNAL  Patent: WO 0231161-A 4 18-Apr-2002;
          BAYER AG (DE)
FEATURES
         source          Location/Qualifiers
                     1..583
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
BASE COUNT  114 a 155 c 190 g 124 t
ORIGIN
Query Match      8.5%; Score 372.2; DB 6; Length 583;
Best Local Similarity 99.2%; Pred. No. 1.9e-86;
Matches 374; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY  4001 AGAGAGGGGACACTGACCTTCTTCCCGAGACTTTTCACTTCTCAGACCGCG 4060
Db  583 AGAGAGGGGACACTGACCTTCTTCCCGAGACTTTTCACTTCTCAGACCGCG 524
QY  4061 GGCATGCCGAGATGCGCATCTGCAACAAATCTGGAACCAAGTGGCCGCAAGA 4120
Db  523 GGCATGCCGAGATGCGCATCTGCAACAAATCTGGAACCAAGTGGCCGCAAGA 464
QY  4121 CTACCTTCAACAACTTCAACCCAGCGGAGCAAACTCAAGTGCCTCTCTGAGAGCC 4180
Db  463 CTACCTTCAACAACTTCAACCCAGCGGAGCAAACTCAAGTGCCTCTCTGAGAGCC 404
QY  4181 CTACCTTCAACAACTTCAACCCAGCGGAGCAAACTCAAGTGCCTCTCTGAGAGCC 4240
Db  403 CTACCTTCAACAACTTCAACCCAGCGGAGCAAACTCAAGTGCCTCTCTGAGAGCC 344
QY  4241 AGGTGCTCTACTGGGCTGTCCCAAGTGGAGCGGAGTGGCTGTGTGGGATCATCG 4300
Db  343 AGGTGCTCTACTGGGCTGTCCCAAGTGGAGCGGAGTGGCTGTGTGGGATCATCG 284
QY  4301 GGAACAGTGTCTGAGAGTGAAGAGAGTGGCCGAGAGGAAGATCTTCAATGAGCCTGC 4360
Db  283 GGAACAGTGTCTGAGAGTGAAGAGAGTGGCCGAGAGGAAGATCTTCAATGAGCCTGC 224
QY  4361 GCACTGTGGCCCTCTAG 4377
Db  223 GCACTGTGGCCCTCTAG 207

```

```

REFERENCE
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
          Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
          Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
          Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
          Diatchenko, L., Marsina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
          Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
          Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Teshigahara, S.,
          Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
          Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.D.,
          McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
          Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hui, X., Gibbs, R.A.,
          Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
          Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
          Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shchegolev, Y.,
          Bouffard, G.G., Blakeley, R.W., Touchman, R.W., Green, E.D.,
          Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
          Butlerfield, Y.S., Krzyzanski, M.I., Skalska, U., Smalins, D.E.,
          Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
          Generation and initial analysis of more than 15,000 full-length
          human and mouse cDNA sequences
          Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL MEDLINE
PUBMED 12388257
REFERENCE 2 (bases 1 to 820)
AUTHORS Strausberg, R.
TITLE    Direct Submission
JOURNAL Submitted (02-JUL-2002) National Institutes of Health, Mammalian
          Gene Collection (MGC), Cancer Genomics Office, National Cancer
          Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
          USA
REMARK
COMMENT NIH-MGC Project URL: http://mgc.ncl.nih.gov
          Contact: MGC help desk
          Email: cgaps-remail.nih.gov
          Tissue Procurement: Gilbert Smith, Ph.D.
          CDNA Library Preparation: Life Technologies, Inc.
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
          DNA Sequencing by: National Institutes of Health Intramural
          Sequencing Center (NISC),
          Gaithersburg, Maryland;
          Web site: http://www.nisc.nih.gov/
          Contact: nisc.mgc@hgtl.nih.gov
          Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
          Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
          Dietrich, N.L., Granice, S., Guan, X., Gupta, J., Haghghi, P.,
          Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
          Maduro, Q.L., Masello, C., Maskeri, B., Mastrlian, S.D., McCloskey, J.C.,
          McDowell, J., Pearson, R., Stantropop, S., Thomas, P.J., Touchman, J.W.,
          Teurigon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
          Young, A., Zhang, L.-H. and Green, E.D.
          Clone distribution: MGC clone distribution information can be found
          through the I.M.A.G.E. Consortium/LNLN at: http://image.llnl.gov
          Series: IRAC Plate: 65 Row: 1 Column: 10
          This clone was selected for full length sequencing because it
          passed the following selection criteria: Hexamer frequency ORF
          analysis.
FEATURES
         source          Location/Qualifiers
                     1..820
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="FVB/N-3"
                     /db_xref="taxon:10090"
                     /clone="MGC:40917 IMAGE:5373588"
                     /tissue_type="Mammary tumor, MMTV-LTR/INT3 model, 5 month
                     old mouse. Taken by biopsy."
                     /clone_1ib="NCI CGAP Mam2"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
                     1..820
                     /gene="MGC40917"
                     /db_xref="LocusID:231108"
                     94..654

```

```

/codon_start=1
/product="hypothetical protein MGCA0917"
/protein_id="AAH33606.1"
/db_xref="GI:21706930"
/db_xref="LOCUSID:231108"
/translation="MOELKLNKNNLOSISELSYWHRYMEREDPAVTVDPPFNTIP
LNREGLDITFSEDCYFSDRGHAELAIALNNMLEPVWMTKSNFNTRKICP
SPERPPLYLIRNSQLLPDKAEPSNMLYVAIPVAAIGLAVGILGVMIMTRVPOOE
EEEDTLPNTSVTVQDAVSEKRLKAGN"
BASE COUNT      266 a      197 c      188 g      169 t
ORIGIN

```

```

Query Match      7.4%; Score 322.2; DB 10; Length 820;
Best Local Similarity 75.1%; Pred. No. 3e-73;
Matches 402; Conservative 0; Mismatches 133; Indels 0; Gaps 0;

```

```

QY 3809 TGGCAGCTCAAACTGCACTTGGCTCAGACCTGCAAGTCCCTGGAGAGCAAG 3868
DB 41 TGTCTTGTAGAAAACTGCAAGTTGCTTAAGACACTCCCAAACTTCACAGCGATGACAG 100
QY 3869 AACTGAAGAAAGTGAAGTGGACCTCCAGATGGCACTCCAGTTTCTCTACTAGGAC 3928
DB 101 AGCTGAAGAACTAACTGGAACCTCCAGAGCGCATCTCGAGCTCTCTATTGGCACC 160
QY 3929 AATACACACAGCGTGAAGACTTTGCGGTGTGTGTCAGACCTTTCTTCCAAAACAACCTCA 3988
DB 161 GGTACATGAGAGCGTGAAGACTTGGCACTGTCGACCTTTCTTCCGAAATACCTTTA 220
QY 3989 CCCCAGTGAACGAGAGAGGAGACACTGACCTCAGCTTTCTTCCGAGACTGTTTCACT 4048
DB 221 TCCCACTGAATGAGCGTGAAGCGCTGACCTCACTTTCTTCTGAAAGACTGTTTCACT 280
QY 4049 TCTCAGACCGCGGAGACTGCGAGATGGGCATGCGACCTGGAACAATGCTGGAACACAG 4108
DB 281 TCTCAGACCGCGTGAAGCTGAGATGGGCATGCGCTTCTTGAATACATCTGGAACACAG 340
QY 4109 TGGCGCCGACAGACTTCTCCACAACTTCAACCAAGCCGACCAAACTCAAGTCCCT 4168
DB 341 TGGCGTGAAGAACTCTCCCAATTAAGCTTCAATATACAAACAACTCAAGTCCCT 400
QY 4169 CTCTGAGAGCCCTTAACTCTTACACCTCGGGAACAGCCGATGCTCTCCAGACAGGCTG 4228
DB 401 CACCTGAAGAGCCCTTCTCTTACACCTCCGGAATAGTCAAGCTTCTTCCAGACAGGCTG 460
QY 4229 AAGAACCCCGGAGGTGCTCTACTGCGGTGTCCAGTGGAGAGGAGGTGCGCTTGTG 4288
DB 461 AAGAACCCCTCAATGCACTTACTGGGCACTGCGCAAGTGGAGCAATAGTGGCCTTGACG 520
QY 4289 TGGGCATCAATCGGAGACAGTGTCTGAGGTGACAGAGAGGTGGCCGAGAGGAGAGA 4343
DB 521 TTGGCATCTTGGAGTGTGTGTGGAGAACTGTGAAACCCGTCACACAGAGAGGA 575

```

```

RESULT 14
AX417822      265 bp      DNA      linear      PAT 18-JUN-2002
LOCUS      AX417822
DEFINITION      Sequence 6 from Patent W00231161.
ACCESSION      AX417822
VERSION      AX417822.1 GI:21522940
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1
Zhu, Z.
Regulation of human phospholipase-like enzyme
Patent: WO 0231161-A 6 18-Apr-2002;
BAYER AG (DE)
FEATURES
Source
1..265
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"

```

```

BASE COUNT      73 a      80 c      62 g      50 t
ORIGIN

```

```

Query Match      5.7%; Score 251; DB 6; Length 265;
Best Local Similarity 100.0%; Pred. No. 1.6e-54;
Matches 251; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3042 CCTTGGACCAATATGCTTGAACCACTTGGACCAAAACAGAGACCTGACCTGAGAGC 3101
DB 15 CTTTGGACCAATATGCTTGAACCACTTGGACCAAAACAGAGACCTTGAACCTGAGAGC 74
QY 3102 AGAGATGCCCATCACTGTCCTCACTGAGATGAGCCCTTCTTGAAGACCTCGGAATAG 3161
DB 75 AGAGATGCCCATCACTGTCCTCACTGAGATGAGCCCTTCTTGAAGACCTCGGAATAG 134
QY 3162 TAACTACAGTACCCCATCAAGCCAGCCATTTGAGAACTGGGGCAGTACTTCTGTATC 3221
DB 135 TAACTACAGTACCCCATCAAGCCAGCCATTTGAGAACTGGGGCAGTACTTCTGTATC 194
QY 3222 AGAGTGAAGGCTTCCAAATGATGTTCCAACTGTCACACAGCTCCGACAGAGACAT 3281
DB 195 AGAGTGAAGGCTTCCAAATGATGTTCCAACTGTCACACAGCTCCGACAGAGACAT 254
QY 3282 CAAAGTGTGG 3292
DB 255 CAAAGTGTGG 265

```

```

RESULT 15
AX417823      446 bp      DNA      linear      PAT 18-JUN-2002
LOCUS      AX417823/c
DEFINITION      Sequence 7 from Patent W00231161.
ACCESSION      AX417823
VERSION      AX417823.1 GI:21522941
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1
Zhu, Z.
Regulation of human phospholipase-like enzyme
Patent: WO 0231161-A 7 18-Apr-2002;
BAYER AG (DE)
FEATURES
Source
1..446
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```

```

BASE COUNT      86 a      122 c      147 g      90 t      1 others
ORIGIN

```

```

Query Match      5.4%; Score 237.4; DB 6; Length 446;
Best Local Similarity 99.2%; Pred. No. 6.4e-51;
Matches 238; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 4138 ACCACAGCCGAGCCAACTCAAGTGCCTCTCTGAGAGCCCTTAACCTTACACCCCTG 4197
DB 446 ACCACAGCCGAGCCAACTCAAGTGCCTCTCTGAGAGCCCTTAACCTTACACCCCTG 387
QY 4198 CGGAACAGCCGATTTGCTCCAGACCAAGGCTGAAGAAAGCCCCGAGGTGCTTACTGGGCT 4257
DB 386 CGGAACAGCCGATTTGCTCCAGACCAAGGCTGAAGAAAGCCCCGAGGTGCTTACTGGGCT 327
QY 4258 GTCCAGTGGCAGCGGAGAGTGGCTTGTGTGTGGCATCATCGGACAGTGTCTGGAGG 4317
DB 326 GTCCAGTGGCAGCGGAGAGTGGCTTGTGTGTGGCATCATCGGACAGTGTCTGGAGG 267
QY 4318 TGCAGAGAGGTGGCCGAGAGGAAAGATCTCCAAATAGCCTGTGGCACTGTGCGCTTTAG 4377
DB 266 TGCAGAGAGGTGGCCGAGAGGAAAGATCTCCAAATAGCCTGTGGCACTGTGCGCTTTAG 207

```

Wed Jan 7 10:03:20 2004

us-10-054-691-1.rge

Page 22

Search completed: January 6, 2004, 12:19:30
Job time : 15568 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 6, 2004, 12:19:45 ; Search time 1300 Seconds
(without alignments)
11653.850 Million cell updates/sec

Title: US-10-054-691-1

Perfect score: 4377
Sequence: 1 atcgggcgtcgccagcagcat.....tgcacactgtgccccttag 4377

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2263443 seqs, 1730637950 residues

Total number of hits satisfying chosen parameters: 4526886

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4377	100.0	4377	US-10-054-691-1	Sequence 1, Appl1
2	970	22.2	2477	US-10-094-749-1391	Sequence 390, App
3	411.2	9.4	572	US-09-764-891-1391	Sequence 1391, App
4	42.4	1.0	1523	US-10-184-644-230	Sequence 290, App
5	42.4	1.0	1523	US-10-184-644-230	Sequence 290, App
6	41.8	1.0	1186	US-10-157-031-268	Sequence 268, App
7	41.8	1.0	2400	US-10-101-510-51	Sequence 51, Appl
8	41.8	1.0	2643	US-10-172-620-15	Sequence 15, Appl
9	41.8	1.0	3633	US-09-725-433-1	Sequence 1, Appl1
10	41.8	1.0	5264	US-09-920-300A-1731	Sequence 1731, App
11	41.8	1.0	5264	US-10-099-928-1731	Sequence 1731, App
12	41.8	1.0	5264	US-10-033-528-1731	Sequence 1731, App
13	41.8	1.0	5532	US-10-007-926A-137	Sequence 137, App
14	41.8	1.0	5532	US-10-101-510-95	Sequence 95, Appl
15	41.8	1.0	5532	US-10-380-931-17	Sequence 17, Appl

C 16	41.8	1.0	10058	10	US-09-974-238-98	Sequence 98, Appl
C 17	41.8	1.0	31056	11	US-09-764-891-6138	Sequence 6138, App
C 18	41.8	1.0	169988	13	US-10-380-931-24	Sequence 24, Appl
C 19	41.8	1.0	197496	10	US-09-877-177-10	Sequence 10, Appl
C 20	41.6	1.0	505	11	US-09-918-995-31558	Sequence 31558, A
C 21	41	0.9	653	15	US-10-184-644-402	Sequence 402, App
C 22	41	0.9	4689	10	US-09-861-289-34	Sequence 34, App
C 23	40.8	0.9	4689	10	US-09-860-846-34	Sequence 34, App
C 24	40.8	0.9	4689	10	US-09-860-846-34	Sequence 34, App
C 25	40.8	0.9	4689	11	US-09-988-384B-34	Sequence 34, App
C 26	40.8	0.9	4689	11	US-09-836-821-34	Sequence 34, App
C 27	40.8	0.9	4689	13	US-10-271-889-34	Sequence 34, App
C 28	40.8	0.9	36778	10	US-09-861-289-5	Sequence 5, Appl1
C 29	40.8	0.9	36778	11	US-09-860-846-5	Sequence 5, Appl1
C 30	40.8	0.9	36778	11	US-09-836-821-5	Sequence 5, Appl1
C 31	40.8	0.9	36778	13	US-10-271-889-48	Sequence 48, Appl
C 32	40.8	0.9	37948	11	US-09-988-384B-5	Sequence 5, Appl1
C 33	40.8	0.9	38506	11	US-09-793-708-19	Sequence 19, Appl
C 34	40.8	0.9	38506	13	US-10-201-365-1	Sequence 1, Appl1
C 35	40.8	0.9	38506	13	US-10-160-539-19	Sequence 19, Appl
C 36	40.2	0.9	877	12	US-10-369-493-30712	Sequence 30712, A
C 37	40.2	0.9	1308	15	US-10-156-761-6349	Sequence 6349, App
C 38	40.2	0.9	9025608	15	US-10-156-761-1	Sequence 1, Appl1
C 39	38.8	0.9	1896	13	US-10-381-779-16	Sequence 16, Appl
C 40	38.8	0.9	225883	15	US-10-175-523-57	Sequence 57, Appl
C 41	38.6	0.9	1428	15	US-10-156-761-2812	Sequence 2812, App
C 42	38	0.9	802	15	US-10-184-644-312	Sequence 312, App
C 43	38	0.9	802	15	US-10-184-644-312	Sequence 312, App
C 44	37.8	0.9	65	13	US-09-908-975-228	Sequence 228, App
C 45	37.8	0.9	1184	13	US-10-140-472-412	Sequence 412, App

ALIGNMENTS

RESULT 1
US-10-054-691-1
; Sequence 1, Application US/10054691
; Publication No. US20020115846A1
; GENERAL INFORMATION:
; APPLICANT: Yu, Xuanchuan
; APPLICANT: Miranda, Maricar
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US0020115846A1 Human Lipase and Polynucleotides Encoding It
; FILE REFERENCE: Lex-0303-USA
; CURRENT APPLICATION NUMBER: US/10/054,691
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: US 60/264,049
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; TYPE: DNA
; LENGTH: 4377
; ORGANISM: homo sapiens
US-10-054-691-1

Query Match	100.0%	Score 4377;	DB 14;	Length 4377;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 4377;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0;
QY	1	ATGGGGCTGGGGCGAGGATTTTCTCTGAGAGCTGCTGCTTCTGGGCGAAGGACC	60	
DB	1	ATGGGGCTGGGGCGAGGATTTTCTCTGAGAGCTGCTGCTTCTGGGCGAAGGACC	60	
QY	61	CCTGAGATCAACCTCTCTAGAAAGATGATGGAAGGAGGAGCTATGGCAGAGACC	120	
DB	61	CCTGAGATCAACCTCTCTAGAAAGATGATGGAAGGAGGAGCTATGGCAGAGACC	120	
QY	121	CTGAAGATTCCTCATTCCTCATGCAACCAATTAATTAAGAGTAATGCTTTAA	180	
DB	121	CTGAAGATTCCTCATTCCTCATGCAACCAATTAATTAAGAGTAATGCTTTAA	180	

Qy 181 TCAGTCTCTGAGCCCTTGATATTAAATTTGGCAGCCATTGGCAATCTGGAA 240
Db 181 TCAGTCTCTGAGCCCTTGATATTAAATTTGGCAGCCATTGGCAATCTGGAA 240
Qy 241 ATTCTTCAGACCCAGGAGCGGCGATCTGGAGAGCAAGACTGGACTGAAAGGCCAG 300
Db 241 ATTCTTCAGACCCAGGAGCGGCGATCTGGAGAGCAAGACTGGACTGAAAGGCCAG 300
Qy 301 CAGGTGACAGGAGGATGACAGCTCTTTCAGACATCATGATATTCAAGCTTCT 360
Db 301 CAGGTGACAGGAGGATGACAGCTCTTTCAGACATCATGATATTCAAGCTTCT 360
Qy 361 GTTCCATGCTGTGTGCACTCTGAGAGAGAGTCAATCCCAAGATGGTGTGAAGAC 420
Db 361 GTTCCATGCTGTGTGCACTCTGAGAGAGAGTCAATCCCAAGATGGTGTGAAGAC 420
Qy 421 TTGTGATTAGGCTCAAGAACTGGTGAAGAACTGAAAGAAAGAACTTGAATTT 480
Db 421 TTGTGATTAGGCTCAAGAACTGGTGAAGAACTGAAAGAAAGAACTTGAATTT 480
Qy 481 CAATTTGACTGGAAGCTCAATATGTTCTTCAATGATGCAAGCCAGTGTACCTGTC 540
Db 481 CAATTTGACTGGAAGCTCAATATGTTCTTCAATGATGCAAGCCAGTGTACCTGTC 540
Qy 541 CCTCTGCTCAACAGAAATGGGCTTTCGCGCGGCGCGTGAATGAGTATGGGGTCTG 600
Db 541 CCTCTGCTCAACAGAAATGGGCTTTCGCGCGGCGCGTGAATGAGTATGGGGTCTG 600
Qy 601 GACTACCTGACAGAGAGTCTCCAGAGCAATTTGTAACTGGTGAAGCTTCTGAGTT 660
Db 601 GACTACCTGACAGAGAGTCTCCAGAGCAATTTGTAACTGGTGAAGCTTCTGAGTT 660
Qy 661 GAGAGGCTCTGCTGATGATCAAGGCACTTGTCAAGCCCTGACAGAGCCCTGTAAT 720
Db 661 GAGAGGCTCTGCTGATGATCAAGGCACTTGTCAAGCCCTGACAGAGCCCTGTAAT 720
Qy 721 TGCTCAGAGAGAGCAACCCGCTGGCCAAAGTGTGATGAGTGTCTTATCAGAAAGC 780
Db 721 TGCTCAGAGAGAGCAACCCGCTGGCCAAAGTGTGATGAGTGTCTTATCAGAAAGC 780
Qy 781 TGGAAAGAGCTCTGCTGCTCCAGAGGTAAGTGAAGAGAGTCTTCAAGCTGTTTC 840
Db 781 TGGAAAGAGCTCTGCTGCTCCAGAGGTAAGTGAAGAGAGTCTTCAAGCTGTTTC 840
Qy 841 CAGGCTTTCTTATGAGAACCCCACTCTACCTGAGAGAGCCCGGCACTCAGAGT 900
Db 841 CAGGCTTTCTTATGAGAACCCCACTCTACCTGAGAGAGCCCGGCACTCAGAGT 900
Qy 901 TCTACACAGCTGAGCTGATCTCTGAAATGAGTGAATGAGAGCAGAGAGAGAAAGAT 960
Db 901 TCTACACAGCTGAGCTGATCTCTGAAATGAGTGAATGAGAGCAGAGAGAGAAAGAT 960
Qy 961 GAGCATTGAGTGTAAACACGAGAGGCAATGAAAGTGTCTCTCAGAGAGAGCCCTAT 1020
Db 961 GAGCATTGAGTGTAAACACGAGAGGCAATGAAAGTGTCTCTCAGAGAGAGCCCTAT 1020
Qy 1021 CTGTTCAGCTACAGAAACAGCACTACCTGACAGAGAGAGAGAGAGAGAGAGAGT 1080
Db 1021 CTGTTCAGCTACAGAAACAGCACTACCTGACAGAGAGAGAGAGAGAGAGAGAGT 1080
Qy 1081 GAGTGAAGT 1140
Db 1081 GAGTGAAGT 1140
Qy 1141 ACCTCAGTTCATGAGCTGAGAGCCGCTGACATCAACGTAATTTGAGCCCTGAGTCT 1200
Db 1141 ACCTCAGTTCATGAGCTGAGAGCCGCTGACATCAACGTAATTTGAGCCCTGAGTCT 1200
Qy 1201 CTCACGAGT 1260
Db 1201 CTCACGAGT 1260
Qy 1261 TACGAGT 1320

Db 1261 TACGAGT 1320
Qy 1321 GCGAATCTCTCGGGAATTTCAACCTTCCCTGAAAGGCTTCTCTGTGGCACTGGGAA 1380
Db 1321 GCGAATCTCTCGGGAATTTCAACCTTCCCTGAAAGGCTTCTCTGTGGCACTGGGAA 1380
Qy 1381 GAAACAGTCTTATGCTTAAACAGAGCTGTGGCAGAGAGAGAGAGAGAGAGAGT 1440
Db 1381 GAAACAGTCTTATGCTTAAACAGAGCTGTGGCAGAGAGAGAGAGAGAGAGAGT 1440
Qy 1441 CCTGTCAGGCGAGAGAGCTGTGACCTGATGAAGATGAACAGAGATACCTTTGAG 1500
Db 1441 CCTGTCAGGCGAGAGAGCTGTGACCTGATGAAGATGAACAGAGATACCTTTGAG 1500
Qy 1501 GAGAGCTGGAAGATTAATACCTGTTTATAGGGGGAATGAGAGCTGTGATTTCTGCAAT 1560
Db 1501 GAGAGCTGGAAGATTAATACCTGTTTATAGGGGGAATGAGAGCTGTGATTTCTGCAAT 1560
Qy 1561 GATCTGATCAATTTCTCCAGAACTTCAAGAGCAATTTGAAAGAGCCCTGAGATC 1620
Db 1561 GATCTGATCAATTTCTCCAGAACTTCAAGAGCAATTTGAAAGAGCCCTGAGATC 1620
Qy 1621 CTCATGCTGAGAGTCTCTGCGGCAATTTGAACTGTGACCTGTGATGATGCTCAAC 1680
Db 1621 CTCATGCTGAGAGTCTCTGCGGCAATTTGAACTGTGACCTGTGATGATGCTCAAC 1680
Qy 1681 CTGAGGAGAGCTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 1740
Db 1681 CTGAGGAGAGCTGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 1740
Qy 1741 TGTCTCTGTCTCTGAAATTTGATGATTAATCAACAGAACTTGTCTCATGATTC 1800
Db 1741 TGTCTCTGTCTCTGAAATTTGATGATTAATCAACAGAACTTGTCTCATGATTC 1800
Qy 1801 AACAGAGATTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 1860
Db 1801 AACAGAGATTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 1860
Qy 1861 GAGATTTTATCTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 1920
Db 1861 GAGATTTTATCTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 1920
Qy 1921 GAAGATGCTGACAACTTTCTTCTGCTCTGAGCTGTTTCACTTCAAGAGAGTCT 1980
Db 1921 GAAGATGCTGACAACTTTCTTCTGCTCTGAGCTGTTTCACTTCAAGAGAGTCT 1980
Qy 1981 CACTCCGAGT 2040
Db 1981 CACTCCGAGT 2040
Qy 2041 ACTGCTATAGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 2100
Db 2041 ACTGCTATAGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 2100
Qy 2101 CTGAGT 2160
Db 2101 CTGAGT 2160
Qy 2161 GCGCTTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 2220
Db 2161 GCGCTTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 2220
Qy 2221 GTGGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 2280
Db 2221 GTGGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 2280
Qy 2281 CTCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 2340
Db 2281 CTCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 2340
Qy 2341 CTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGT 2400

Db 2341 CTGGAGATGTGACCACTTACTATATCTTCGGGAGTTTAAACAGAACTCAGAGC 2400
 Qy 2401 TAGCCGTGGGACGGGTGATGCAATGACAGCAATGATCTTCCATCAAGCGTTCC 2460
 Db 2401 TAGCCGTGGGACGGGTGATGCAATGACAGCAATGATCTTCCATCAAGCGTTCC 2460
 Qy 2461 GGAGCAAGGCTGAGATCTTATGAGCCAAAGTCCAACTCTGATGAGAAATGAAAT 2520
 Db 2461 GGAGCAAGGCTGAGATCTTATGAGCCAAAGTCCAACTCTGATGAGAAATGAAAT 2520
 Qy 2521 GATCATAGATTAATTTCCATGAAAGCTGAGAGCTCATCAAGTCTGATGAGAGC 2580
 Db 2521 GATCATAGATTAATTTCCATGAAAGCTGAGAGCTCATCAAGTCTGATGAGAGC 2580
 Qy 2581 GATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2640
 Db 2581 GATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2640
 Qy 2641 CTCGGCAATGCTTGGACGTCCTGCAATGAGAGGTGCGAGATCTGATCACTCTG 2700
 Db 2641 CTCGGCAATGCTTGGACGTCCTGCAATGAGAGGTGCGAGATCTGATCACTCTG 2700
 Qy 2701 GACTTCTGTAACCCCACTATCATGCGGCAAGTGTCTTGGGAAACCAAGTCCCA 2760
 Db 2701 GACTTCTGTAACCCCACTATCATGCGGCAAGTGTCTTGGGAAACCAAGTCCCA 2760
 Qy 2761 GTGACAGAGGCGAGGCTTGTGTAATGATGATGATGATGATGATGATGATGATGAT 2820
 Db 2761 GTGACAGAGGCGAGGCTTGTGTAATGATGATGATGATGATGATGATGATGATGAT 2820
 Qy 2821 CTAGCCAGGCTGAGAGGCTTCAAGCCAGCTTACCGAGAGCATGCCGAGCTGAGG 2880
 Db 2821 CTAGCCAGGCTGAGAGGCTTCAAGCCAGCTTACCGAGAGCATGCCGAGCTGAGG 2880
 Qy 2881 TCAGGCGCTGATACACGACGAGAGGATCTTGTGTGCTGAGAGGCTTCTTCCAGAC 2940
 Db 2881 TCAGGCGCTGATACACGACGAGAGGATCTTGTGTGCTGAGAGGCTTCTTCCAGAC 2940
 Qy 2941 ATCAGGCTTCTTCTGCGGAGATGAGGCTCCAGATACGTCCTTCTTCCAGAGTGC 3000
 Db 2941 ATCAGGCTTCTTCTGCGGAGATGAGGCTCCAGATACGTCCTTCTTCCAGAGTGC 3000
 Qy 3001 ATCCACCCAAATGAGAAATTCCTCCAGCTGAGAGGCTTGGAGCAATATGCTT 3060
 Db 3001 ATCCACCCAAATGAGAAATTCCTCCAGCTGAGAGGCTTGGAGCAATATGCTT 3060
 Qy 3061 GAAACAATTGAGAGCAAGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGG 3120
 Db 3061 GAAACAATTGAGAGCAAGAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTGAGAGG 3120
 Qy 3121 CCCACTCAGATGAGGCTTCTGAGAGGCTTCTGAGAGGCTTCTGAGAGGCTTCTGAG 3180
 Db 3121 CCCACTCAGATGAGGCTTCTGAGAGGCTTCTGAGAGGCTTCTGAGAGGCTTCTGAG 3180
 Qy 3181 AAGCCAGGCTTGAAGAGGCTGAGAGGCTGAGAGGCTGAGAGGCTTCTGAGAGG 3240
 Db 3181 AAGCCAGGCTTGAAGAGGCTGAGAGGCTGAGAGGCTTCTGAGAGGCTTCTGAGAGG 3240
 Qy 3241 AGTGTTCACACCTTGTCCACAGCTCCGACAGAGAGATCAAGAGTGTGTGCGCTG 3300
 Db 3241 AGTGTTCACACCTTGTCCACAGCTCCGACAGAGAGATCAAGAGTGTGTGCGCTG 3300
 Qy 3301 GGTGATCTTGTGATGATGAGAGGCTGAGAGGCTGAGAGGCTTCTGAGAGGCTTCTGAG 3360
 Db 3301 GGTGATCTTGTGATGATGAGAGGCTGAGAGGCTTCTGAGAGGCTTCTGAGAGGCTTCTGAG 3360
 Qy 3361 TCTTGGAGGAGATCTTCTTGGAGATGAGAGGAGTGTGAGATCTTCAACACACA 3420
 Db 3361 TCTTGGAGGAGATCTTCTTGGAGATGAGAGGAGTGTGAGATCTTCAACACACA 3420
 Qy 3421 CTGCCCCAATTTGAGAGAGTCAACCTTACCTCTTGGCTTCTTACAGACAGCTG 3480
 Db 3421 CTGCCCCAATTTGAGAGAGTCAACCTTACCTCTTGGCTTCTTACAGACAGCTG 3480

Qy 3481 GAGGGGACAGCAGAGATTAATGTCAGCGGAAAGGGGCGAGCTAGGAGCATGCCAGCC 3540
 Db 3481 GAGGGGACAGCAGAGATTAATGTCAGCGGAAAGGGGCGAGCTAGGAGCATGCCAGCC 3540
 Qy 3541 CAGGCTTGGAGCTTGTAGAGCGAATGAAACAGCCCCGACATCAACCTGAGAAAGAC 3600
 Db 3541 CAGGCTTGGAGCTTGTAGAGCGAATGAAACAGCCCCGACATCAACCTGAGAAAGAC 3600
 Qy 3601 TGGAGCTGTGATCACTTCTTCAATGGGGTCAACGACTTGTGTATTAATGTGAATCCG 3660
 Db 3601 TGGAGCTGTGATCACTTCTTCAATGGGGTCAACGACTTGTGTATTAATGTGAATCCG 3660
 Qy 3661 GAGGCGCTTGGGCGAGGAAATGTTTCAAGCATCAACAGGCGCTGAGCATCTCT 3720
 Db 3661 GAGGCGCTTGGGCGAGGAAATGTTTCAAGCATCAACAGGCGCTTGGAGCATCTCT 3720
 Qy 3721 GAGAGCTCCCAAGGCTTTCGTCAACGTGTGAGAGTCAATGAGCTGTGAGCTGTAC 3780
 Db 3721 GAGAGCTCCCAAGGCTTTCGTCAACGTGTGAGAGTCAATGAGCTGTGAGCTGTAC 3780
 Qy 3781 CAGGCGCAAGGCGGAAATGTGCTATGCTGAGCTGAGAGCTCAAGCAATGCTGACAG 3840
 Db 3781 CAGGCGCAAGGCGGAAATGTGCTATGCTGAGCTGAGAGCTCAAGCAATGCTGACAG 3840
 Qy 3841 CACTGCAAGAGCTCCCTGAGAGAGCAAGAACTGAAAGTGAACCTGCAAGCAT 3900
 Db 3841 CACTGCAAGAGCTCCCTGAGAGAGCAAGAACTGAAAGTGAACCTGCAAGCAT 3900
 Qy 3901 GGCATCTCCAGATTTCTCTACTAGGCAACATACACAGCGTGAAGACTTTCGATG 3960
 Db 3901 GGCATCTCCAGATTTCTCTACTAGGCAACATACACAGCGTGAAGACTTTCGATG 3960
 Qy 3961 GTGACGCTTCTTCTTCCAAACACATCAACCTCACTGAAAGAGGAGCACTGACCTC 4020
 Db 3961 GTGACGCTTCTTCTTCCAAACACATCAACCTCACTGAAAGAGGAGCACTGACCTC 4020
 Qy 4021 ACCCTTCTCCGAGAGCTGTTTCACTTCTGAGAGCGGAGGATGCGAGATGAGCATC 4080
 Db 4021 ACCCTTCTCCGAGAGCTGTTTCACTTCTGAGAGCGGAGGATGCGAGATGAGCATC 4080
 Qy 4081 GCACTGTGAAACAACTGCTGAGACAGTGTGAGGCGGAGAGTCACTTCAACATTCACC 4140
 Db 4081 GCACTGTGAAACAACTGCTGAGACAGTGTGAGGCGGAGAGTCACTTCAACATTCACC 4140
 Qy 4141 CACAGCGAGCCAACTCAAGTGCCTCTCTGAGAGGCTTCAACCTTCAACCTTGC 4200
 Db 4141 CACAGCGAGCCAACTCAAGTGCCTCTCTGAGAGGCTTCAACCTTCAACCTTGC 4200
 Qy 4201 AACAGCGATGCTTCCCAACAGAGGCTGAAAGAGCCCGAGAGTGTCTATGAGGCTG 4260
 Db 4201 AACAGCGATGCTTCCCAACAGAGGCTGAAAGAGCCCGAGAGTGTCTATGAGGCTG 4260
 Qy 4261 CCAAGTGGAGCGGAGTGGCTTGTGTGTGAGGATCATCGGAGAGTGTGTGAGGAGTGC 4320
 Db 4261 CCAAGTGGAGCGGAGTGGCTTGTGTGTGAGGATCATCGGAGAGTGTGTGAGGAGTGC 4320
 Qy 4321 AGGAGAGGTGGCGGAGGAGAGATCTTCAATGAGCTTGCAGCTGTGCGCTCTAG 4377
 Db 4321 AGGAGAGGTGGCGGAGGAGAGATCTTCAATGAGCTTGCAGCTGTGCGCTCTAG 4377

RESULT 2
 US-10-094-749-390
 ; Sequence 390, Application US/10094749
 ; Publication No. US20030219741A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ISOGAI, TAKAO
 ; APPLICANT: SUGIYAMA, TOMOYASU
 ; APPLICANT: OTSUKI, TETSUJI
 ; APPLICANT: WAKAMATSU, AI
 ; APPLICANT: SATO, HIROYUKI
 ; APPLICANT: ISHII, SHIZUKO

APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHICO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOKU
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
FILE REFERENCE: 084335/0160
CURRENT APPLICATION NUMBER: US/10/094,749
CURRENT FILING DATE: 2002-03-12
PRIOR APPLICATION NUMBER: 60/350,435
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: JP 2001-328381
PRIOR FILING DATE: 2001-09-14
NUMBER OF SEQ ID NOS: 3381
SOFTWARE: Patent Ver. 2.1
SEQ ID NO 390
LENGTH: 2477
TYPE: DNA
ORGANISM: Homo sapiens
US-10-094-749-390

Query Match 22.4%; Score 970; DB 13; Length 2477;
Best Local Similarity 99.5%; Pred. No. 7.8e-259;
Matches 973; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1279 AGCGTCGGCGGAGTGAAGATCGGACCGTTACCACTCGGGGAACTCTCCGGAA 1338
DB 6 AGCGTCGGCGGAGTGAAGATCGGACCGTTACCACTCGGGGAACTCTCCGGAA 65
QY 1339 TTCACACCTTCCCGAGAGGGCTCTCTGTTGGCACTGGGAAAGAACCACTCTATGCC 1398
DB 66 TTCACACCTTCCCGAGAGGGCTCTCTGTTGGCACTGGGAAAGAACCACTCTATGCC 125
QY 1399 TTCTTAACAGAGCTGTGGCAGAGGCGAGCTGAGGATCTACCTGTCCAGGCCAGAGG 1458
DB 126 TTCTTAACAGAGCTGTGGCAGAGGCGAGCTGAGGATCTACCTGTCCAGGCCAGAGG 185
QY 1459 CTGTGTGACCTGATGAGGATGACAGAGATACCTTCAAGAAAGCTGGAATATATA 1518
DB 186 CTGTGTGACCTGATGAGGATGACAGAGATACCTTCAAGAAAGCTGGAATATATA 245
QY 1519 ACCCTGTTATAGGCGGCAATGACCTGTGATTTCTGCAATGATCTGTCCACTATCT 1578
DB 246 ACCCTGTTATAGGCGGCAATGACCTGTGATTTCTGCAATGATCTGTCCACTATCT 305
QY 1579 CCCGAGAACTTCAAGACAACTTGAAGAGCCCTGGAATCTCTCATGTGAGTTCT 1638
DB 306 CCCGAGAACTTCAAGACAACTTGAAGAGCCCTGGAATCTCTCATGTGAGTTCT 365
QY 1639 CGGGCAATTTGGAACCTGTGAGAGCGGTCTGATGATCTCACTGAGAGAGCTTACAG 1698
DB 366 CGGGCAATTTGGAACCTGTGAGAGCGGTCTGATGATCTCACTGAGAGAGCTTACAG 425
QY 1699 GAGAAAAAGTCTCTGCGCAAGATGATCTCTGAGTCTCTGATCTCTGATCTCTGAG 1758
DB 426 GAGAAAAAGTCTCTGCGCAAGATGATCTCTGAGTCTCTGATCTCTGATCTCTGAG 485
QY 1759 TTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1818
DB 486 TTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 545
QY 1819 AAGACCAACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1878
DB 546 AAGACCAACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 605
QY 1879 GTGAGAGCGTTCTTTGAAAAAGTGGACATGCAAGAACCTCGAAGAGATTGCCGACAC 1938

DB 606 GTGAGAGCGTTCTTTGAAAAAGTGGACATGCAAGAACCTCGAAGAGATTGCCGACAC 665
QY 1939 TCTTTCTGCTCTGATGCTGTTTCCATCTTCAAGAGCAATCTTCCGAGACCACT 1998
DB 666 TCTTTCTGCTCTGATGCTGTTTCCATCTTCAAGAGCAATCTTCCGAGACCACT 725
QY 1999 GCTCTGGAACCAATATGCTGAGAGCGGTGAGGAGAGAGAGAGAGAGAGAGAGAG 2058
DB 726 GCTCTGGAACCAATATGCTGAGAGCGGTGAGGAGAGAGAGAGAGAGAGAGAGAG 785
QY 2059 AACAGATCAATATCAATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2118
DB 786 AACAGATCAATATCAATGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 845
QY 2119 AGCATGAGGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2178
DB 846 AGCATGAGGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 905
QY 2179 CCTACCTGATGATGAG 2238
DB 906 CCTACCTGATGATGAG 965
QY 2239 TCTCTGACCGCTGCAAT 2256
DB 966 TCTCTGACCGCTGCAAT 983

RESULT 3
US-09-764-891-1391
Sequence 1391, Application US/09764891
Publication No. US20030077808A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC006
CURRENT APPLICATION NUMBER: US/09/764,891
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PAM or file wrapper
NUMBER OF SEQ ID NOS: 10231
SOFTWARE: Patent Ver. 2.0
SEQ ID NO 1391
LENGTH: 572
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (375)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (529)
OTHER INFORMATION: n equals a,t,g, or c
US-09-764-891-1391
Query Match 9.4%; Score 411.2; DB 11; Length 572;
Best Local Similarity 96.7%; Pred. No. 2.8e-120;
Matches 492; Conservative 2; Mismatches 6; Indels 9; Gaps 7;
QY 3042 CCTTTGAGCAATATGCTTGAACCACTTGAAGCAAAAGAGAGAGAGAGAGAGAGAG 3101
DB 15 CCTTTGAGCAATATGCTTGAACCACTTGAAGCAAAAGAGAGAGAGAGAGAGAGAG 74
QY 3102 AGAGATGCCATCACTGTTCCACTGAGATGAGCCCTTCTTGAAGAGAGAGAGAG 3161
DB 75 AGAGATGCCATCACTGTTCCACTGAGATGAGCCCTTCTTGAAGAGAGAGAGAGAG 134
QY 3162 TAACCTAGTACCACTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3221
DB 135 TAACCTAGTACCACTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 194
QY 3222 AGAGTGAAGAGCTTCAATAGTGTTCACCTCTGTCCAGAGAGAGAGAGAGAGAG 3281
DB 195 AGAGTGAAGAGCTTCAATAGTGTTCACCTCTGTCCAGAGAGAGAGAGAGAGAG 254


```

APPLICANT: Krutkovskaya, L. L.
TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
FILE REFERENCE: 2760-103
CURRENT APPLICATION NUMBER: US/10/157,031
CURRENT FILING DATE: 2002-05-30
NUMBER OF SEQ ID NOS: 415
SOFTWARE: PatentIn version 3.1
SEQ ID NO 268
LENGTH: 1186
TYPE: DNA
ORGANISM: Homo sapiens
US-10-157-031-268

Query Match      1.0%; Score 42; DB:15; Length 1186;
Beet Local Similarity 50.5%; Pred. No. 0.076;
Matches 102; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY      2794      CTGACCCGCGCGGAACTCCCAAGAGCTACCCAGAGCTTGAGAGCCCTTACCGAGCCTAC 2853
Db      568      CCGCCGCGTGCCTCGGCGGCTCGCCGCGCGGCTCATCTGTCAGAGCCTCCACCGCGAGCTG 627

QY      2854      CGAGCAGCATGCGCGAGCTGTGGGTACAGCCCGCTATGACACGAGAGGACTTTCT 2913
Db      628      CTGGCGTGTCTGGGCGACTGTGTGCGGACGCTCTATGCGCTCAAGAGCTGAGCCG 687

QY      2914      GTGTGCTGCAAGCCCTTCTTCCAGAACTCCAGCTCCTGTGTCCTGAGCGATGGGCTCCCA 2973
Db      688      GGCAGGCTGTGGGCTGTCTCCGGGTGTGACCGCTGTGCTGCTGTGAGGGCTGGCAA 747

QY      2974      GATACGTCTCTTCTTGCCCCAG 2995
Db      748      GACCAGGCTTCAATTAGCCTTG 769

```

```

RESULT 7
US-10-101-510-51/c
; Sequence 51, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: MANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101.510
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
; LENGTH: 2400
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-51

Query Match          1.0%; Score 41.8; DB 13; Length 2400;
Best Local Similarity 61.5%; Pred. No. 0.14;
Matches    67; Conservative   0; Mismatches    42; Indels      0; Gaps    0;

QY      4206  CCAGTTCTCCCAACGAGGTGAAGAAGCCCCCGAGGTGCTCACAAGGGCTGTCGCCAGT 4265
           ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       1408  CTGGTACTTCCACAGACGAGGTGTGTTTTCTCCCAATGACTCTGC CGGSCAGTCTTGA 1348

QY      4266  GGACCGCGAGTCGGCTTGTTGTTGGGCATCATCGGACAGTAGTGTGG 4314
           ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db       1348  CGCAGTGGGGGCCGTCAATGTAGTGGGCACTGGATTACAGTTGTCTGG 1300

RESULT 8
US-10-172-620-15/c
; Sequence 15, Application US/10172620
; Publication No. US2003005395A1
; GENERAL INFORMATION:
```

APPLICANT: Hung, Mien-Chie
APPLICANT: Lin, Shiao-Yih
TITLE OF INVENTION: Methods and Compositions for Inhibiting EGF Receptor
FILE REFERENCE: UTSC:720US
CURRENT APPLICATION NUMBER: US/10/172,620
CURRENT FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: US 60/298,579
PRIOR FILING DATE: 2001-06-15
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.1
SEQ ID NO 15
LENGTH: 2643
TYPE: DNA
ORGANISM: Human
US-10-172-620-15

Query Match 1.0%; Score 41.8; DB 15; Length 2643;
Best Local Similarity 61.5%; Pred. No. 0.15;
Matches 67; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 4206 CCGATTGCTCCAGACAGGCTGAGAGCCCGGAGGTGCTCTACTGGGCTGTCCAGT 4265
DB 1993 CTGCGTACTTCCAGACAGGAGGTGTGTTTCTCCATGACTCTGCGGAGGCTTGA 1934
QY 4266 GGCAGCGGAGTGGCTGTGTGGTCATCATCGGAGACAGTGTCTGG 4314
DB 1993 CGCAGTGGGGCCGTCAATGATGAGGACACTGATACAGTTGTCTGG 1885

RESULT 9
US-09-725-433-1/c
Sequence 1, Application US/09725433
Patent No. US20020068362A1
GENERAL INFORMATION:
APPLICANT: No. US20020068362A1artis AG
TITLE OF INVENTION: Increased transgene expression in retroviral vectors having a sca
FILE REFERENCE: 4-30921B/SYS
CURRENT APPLICATION NUMBER: US/09/725,433
CURRENT FILING DATE: 2000-11-29
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 3633
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(3633)
OTHER INFORMATION:
US-09-725-433-1

Query Match 1.0%; Score 41.8; DB 9; Length 3633;
Best Local Similarity 61.5%; Pred. No. 0.18;
Matches 67; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 4206 CCGATTGCTCCAGACAGGCTGAGAGCCCGGAGGTGCTCTACTGGGCTGTCCAGT 4265
DB 1834 CTGCGTACTTCCAGACAGGAGGTGTGTTTCTCCATGACTCTGCGGAGGCTTGA 1775
QY 4266 GGCAGCGGAGTGGCTGTGTGGTCATCATCGGAGACAGTGTCTGG 4314
DB 1774 CGCAGTGGGGCCGTCAATGATGAGGACACTGATACAGTTGTCTGG 1726

RESULT 10
US-09-920-300A-1731/c
Sequence 1731, Application US/09920300A
Patent No. US20020136728A1
GENERAL INFORMATION:
APPLICANT: King, Gordon E.
APPLICANT: Meagher, Madeleine Joy
APPLICANT: Xu, Jiangchun

APPLICANT: Secrist, Heather
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121,547
CURRENT APPLICATION NUMBER: US/09/920,300A
CURRENT FILING DATE: 2001-07-31
NUMBER OF SEQ ID NOS: 1789
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1731
LENGTH: 5264
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2497
OTHER INFORMATION: n = A,T,C or G
US-09-920-300A-1731

Query Match 1.0%; Score 41.8; DB 10; Length 5264;
Best Local Similarity 61.5%; Pred. No. 0.22;
Matches 67; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 4206 CCGATTGCTCCAGACAGGCTGAGAGCCCGGAGGTGCTCTACTGGGCTGTCCAGT 4265
DB 1753 CTGCGTACTTCCAGACAGGAGGTGTGTTTCTCCATGACTCTGCGGAGGCTTGA 1694
QY 4266 GGCAGCGGAGTGGCTGTGTGGTCATCATCGGAGACAGTGTCTGG 4314
DB 1693 CGCAGTGGGGCCGTCAATGATGAGGACACTGATACAGTTGTCTGG 1645

RESULT 11
US-10-099-926-1731/c
Sequence 1731, Application US/10099926
Publication No. US2003016064A1
GENERAL INFORMATION:
APPLICANT: King, Gordon E.
APPLICANT: Meagher, Madeleine Joy
APPLICANT: Xu, Jiangchun
APPLICANT: Secrist, Heather
APPLICANT: Jiang, Yugu
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121,547C2
CURRENT APPLICATION NUMBER: US/10/099,926
CURRENT FILING DATE: 2002-03-17
NUMBER OF SEQ ID NOS: 1982
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1731
LENGTH: 5264
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2497
OTHER INFORMATION: n = A,T,C or G
US-10-099-926-1731

Query Match 1.0%; Score 41.8; DB 13; Length 5264;
Best Local Similarity 61.5%; Pred. No. 0.22;
Matches 67; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 4206 CCGATTGCTCCAGACAGGCTGAGAGCCCGGAGGTGCTCTACTGGGCTGTCCAGT 4265
DB 1753 CTGCGTACTTCCAGACAGGAGGTGTGTTTCTCCATGACTCTGCGGAGGCTTGA 1694
QY 4266 GGCAGCGGAGTGGCTGTGTGGTCATCATCGGAGACAGTGTCTGG 4314
DB 1693 CGCAGTGGGGCCGTCAATGATGAGGACACTGATACAGTTGTCTGG 1645

RESULT 12
US-10-033-528-1731/c

; Sequence 1731, Application US/10033528
; Publication No. US2002013197A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547C1
; CURRENT APPLICATION NUMBER: US/10/033.528
; CURRENT FILING DATE: 2001-12-26
; NUMBER OF SEQ ID NOS: 1896
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1731
; LENGTH: 5264
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2497
; OTHER INFORMATION: n = A,T,C or G
US-10-033-528-1731

Query Match 1.0%; Score 41.8; DB 14; Length 5264;
Best Local Similarity 61.5%; Pred. No. 0.22;
Matches 67; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 4206 CCGATTGCTCCACAGACGAGGCTGAAGAGCCCGAGAGTGTCTACTGGGCTGCCAGT 4265
Db 1753 CTGGTACTTCCAGACAGGAGGTGTGTTTCTCCATGACTCCGCGGACAGGTCTTGA 1694

Qy 4266 GGCAGCGGAGTGGCGCTTGTGTGGGCATCATCGGAGGAGTGTCTGG 4314
Db 1693 CGCAGTGGGCGCGTCAATGATGAGGACACTGATACAGTGTCTGG 1645

RESULT 13

US-10-007-926A-137/c

; Sequence 137, Application US/10007926A
; Publication No. US20030143539A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, FRANCOIS
; APPLICANT: HOUJGATTE, REMI
; APPLICANT: BIRNBAUM, DANIEL
; APPLICANT: NGUYEN, CATHERINE
; APPLICANT: VIENS, PATRICE
; APPLICANT: FERT, VINCENT
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
; TITLE OF INVENTION: USING ARRAYS OF CANDIDATE GENES
; FILE REFERENCE: 1546-R-00
; CURRENT APPLICATION NUMBER: US/10/007.926A
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: 60/254,090
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 468
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 137
; LENGTH: 5532
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: epidermal growth factor receptor (avian
; OTHER INFORMATION: erythroid-leukemia viral (v-erb-b) oncogene
; OTHER INFORMATION: homology) (EGFR) gene.
US-10-007-926A-137

Query Match 1.0%; Score 41.8; DB 13; Length 5532;
Best Local Similarity 61.5%; Pred. No. 0.23;
Matches 67; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 4206 CCGATTGCTCCACAGACGAGGCTGAAGAGCCCGAGAGTGTCTACTGGGCTGCCAGT 4265
Db 2020 CTGGTACTTCCAGACAGGAGGTGTGTTTCTCCATGACTCTGCGGACAGGTCTTGA 1961

Db 2020 CTGGTACTTCCAGACAGGAGGTGTGTTTCTCCATGACTCTGCGGACAGGTCTTGA 1961
Qy 4266 GGCAGCGGAGTGGCGCTTGTGTGGGCATCATCGGAGGAGTGTCTGG 4314
Db 1960 CGCAGTGGGCGCGTCAATGATGAGGACACTGATACAGTGTCTGG 1912

RESULT 14

US-10-101-510-95/c

; Sequence 95, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: MAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101.510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 95
; LENGTH: 5532
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-95

Query Match 1.0%; Score 41.8; DB 13; Length 5532;
Best Local Similarity 61.5%; Pred. No. 0.23;
Matches 67; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 4206 CCGATTGCTCCACAGACGAGGCTGAAGAGCCCGAGAGTGTCTACTGGGCTGCCAGT 4265
Db 2020 CTGGTACTTCCAGACAGGAGGTGTGTTTCTCCATGACTCTGCGGACAGGTCTTGA 1961

Qy 4266 GGCAGCGGAGTGGCGCTTGTGTGGGCATCATCGGAGGAGTGTCTGG 4314
Db 1960 CGCAGTGGGCGCGTCAATGATGAGGACACTGATACAGTGTCTGG 1912

RESULT 15

US-10-380-931-17/c

; Sequence 17, Application US/10380931
; Publication No. US20030215944A1
; GENERAL INFORMATION:
; APPLICANT: Isis Pharmaceuticals, Inc.
; APPLICANT: C. Frank Bennett
; APPLICANT: Jacqueline Wyatt
; APPLICANT: Susan M. Freier
; TITLE OF INVENTION: OLIGONUCLEOTIDE INHIBITION OF HER-1 EXPRESSION
; FILE REFERENCE: RTSP-0187
; CURRENT APPLICATION NUMBER: US/10/380.931
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: 09/676,610
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 182
; SEQ ID NO 17
; LENGTH: 5532
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (187)...(3819)
US-10-380-931-17

Query Match 1.0%; Score 41.8; DB 13; Length 5532;
Best Local Similarity 61.5%; Pred. No. 0.23;
Matches 67; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 4206 CCGATTGCTCCACAGACGAGGCTGAAGAGCCCGAGAGTGTCTACTGGGCTGCCAGT 4265
Db 2020 CTGGTACTTCCAGACAGGAGGTGTGTTTCTCCATGACTCTGCGGACAGGTCTTGA 1961

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 6, 2004, 06:59:00 ; Search time 236 Seconds
(without alignments)
8186.161 Million cell updates/sec

Title: US-10-054-691-1

Perfect score: 4377
Sequence: 1 atggggctgcggccagcat.....tgcacactgtgcccctctag 4377

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	41.8	1.0	5532	2 US-08-475-035-3	Sequence 3, Appl1
C 2	41.8	1.0	5532	4 US-09-676-6108-17	Sequence 17, Appl1
C 3	41.8	1.0	169998	4 US-09-676-6108-24	Sequence 24, Appl1
C 4	41.8	1.0	197486	4 US-09-877-177A-10	Sequence 10, Appl1
5	40.8	0.9	4689	3 US-09-105-537-34	Sequence 34, Appl1
6	40.8	0.9	36778	3 US-09-105-537-5	Sequence 5, Appl1
7	40.8	0.9	38506	3 US-09-320-878-19	Sequence 19, Appl1
8	40.8	0.9	38506	4 US-09-141-908-1	Sequence 1, Appl1
9	40.8	0.9	38506	4 US-09-657-440-19	Sequence 19, Appl1
10	40.8	0.9	1465	4 US-09-338-671-1	Sequence 1, Appl1
11	39.6	0.9	800	5 PCT-US95-04801-4	Sequence 4, Appl1
C 12	38	0.9	7218	1 US-08-232-463-14	Sequence 14, Appl1
C 13	37.6	0.9	513	4 US-09-252-991A-5791	Sequence 5791, Ap
C 14	37.6	0.9	762	4 US-09-252-991A-5809	Sequence 5809, Ap
C 15	37.6	0.9	1077	4 US-09-252-991A-5839	Sequence 5839, Ap
16	37.6	0.9	2007	4 US-09-252-991A-5770	Sequence 5770, Ap
17	37.2	0.8	840	4 US-09-252-991A-1634	Sequence 1634, Ap
18	37.2	0.8	1335	4 US-09-252-991A-973	Sequence 973, Ap
19	37.2	0.8	1335	4 US-09-252-991A-1685	Sequence 1685, Ap
C 20	37.2	0.8	1653	4 US-09-252-991A-1111	Sequence 1111, Ap
C 21	37.2	0.8	4563	4 US-09-252-991A-930	Sequence 930, App
22	37.2	0.8	1413	4 US-09-252-991A-6071	Sequence 6071, Ap
C 23	37	0.8	1431	4 US-09-252-991A-6232	Sequence 6232, Ap
C 24	37	0.8	1506	4 US-09-252-991A-6228	Sequence 6228, Ap
C 25	37	0.8	2005	4 US-09-635-872A-18	Sequence 18, Appl
C 26	37	0.8	2005	4 US-09-636-077A-18	Sequence 18, Appl
27	36.8	0.8	1555	2 US-08-696-376-1	Sequence 1, Appl1

28	36.8	0.8	3150	3 US-08-943-768-1	Sequence 1, Appl1
29	36.8	0.8	3150	4 US-09-865-960-1	Sequence 1, Appl1
30	36.6	0.8	2502	1 US-08-073-384C-7	Sequence 7, Appl1
31	36.6	0.8	2502	1 US-08-254-359A-7	Sequence 7, Appl1
32	36.6	0.8	2502	1 US-08-483-043-7	Sequence 7, Appl1
33	36.6	0.8	2502	1 US-08-481-238-7	Sequence 7, Appl1
34	36.6	0.8	2502	2 US-08-471-066B-7	Sequence 7, Appl1
35	36.6	0.8	2502	2 US-08-484-956-7	Sequence 7, Appl1
36	36.6	0.8	2502	2 US-08-757-653-7	Sequence 7, Appl1
37	36.6	0.8	2502	2 US-08-599-491-7	Sequence 7, Appl1
38	36.6	0.8	2502	2 US-08-756-386-7	Sequence 7, Appl1
39	36.6	0.8	2502	2 US-08-823-516-7	Sequence 7, Appl1
40	36.6	0.8	2502	3 US-08-682-853A-7	Sequence 7, Appl1
41	36.6	0.8	2502	3 US-08-759-038-7	Sequence 7, Appl1
42	36.6	0.8	2502	3 US-08-758-314-7	Sequence 7, Appl1
43	36.6	0.8	2502	4 US-09-350-309-7	Sequence 7, Appl1
44	36.6	0.8	2502	4 US-08-520-946-7	Sequence 7, Appl1
45	36.6	0.8	2502	4 US-09-684-938-7	Sequence 7, Appl1

ALIGNMENTS

RESULT 1
US-08-475-035-3/C
Sequence 3, Application US/08475035

Patent No. 5985553

GENERAL INFORMATION:

APPLICANT: KING, C. R.

APPLICANT: KRAUS, MATTHIAS H.

TITLE OF INVENTION: HUMAN GENE RELATED TO BUT DISTINCT FROM

TITLE OF INVENTION: BGF RECEPTOR GENE

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESSES:

ADDRESSEE: NEEDLE & ROSENBERG, P.C.

STREET: Suite 1200, 127 Peachtree Street

CITY: Atlanta

STATE: Georgia

COUNTRY: USA

ZIP: 30303

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/475,035

FILING DATE: 7 Jun 1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Perryman, David G.

REGISTRATION NUMBER: 33,438

REFERENCE/DOCKET NUMBER: 1414.656

TELECOMMUNICATION INFORMATION:

TELEPHONE: 404/688-0770

TELEFAX: 404/688-9880

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 5532 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 187..3816

Query Match 1.0%; Score 41.8; DB 2; Length 5532;

Best Local Similarity 61.5%; Pred. No. 0.27;

Matches 67; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 4206 CCATTGCTCCAGACAGGCTGAGAGCCCGAGTCTTACTGGGCTGTCCAGT 4265
Db 147498 CTCGACTCTCCAGACAGGCTGTGTTTCTCCAGTACTCTGCGCGGAGGTCTTGA 147439
QY 4266 GCGAGGGAGTGGCTTGTGTGATCATTCGGGACAGTGGTCTGG 4314
Db 147438 CGCAGTGGGGGCGTCAATGTAGTGGACACATGGATACAGTTGTCTGG 147390

RESULT 4

US-09-877-177A-10/c
; Sequence 10, Application US/09877177A
; Patent No. 6582919
; GENERAL INFORMATION:
; APPLICANT: K. Danenberg
; TITLE OF INVENTION: Method of determining Epidermal Growth
; TITLE OF INVENTION: Factor Receptor and HER2-Neu Gene Expression
; FILE REFERENCE: 11220/120
; CURRENT APPLICATION NUMBER: US/09/877,177A
; CURRENT FILING DATE: 2001-06-11
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 197496
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-877-177A-10

Query Match 1.0%; Score 41.8; DB 4; Length 197496;
Best Local Similarity 61.5%; Pred. No. 2.3;
Matches 67; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 4206 CCATTGCTCCAGACAGGCTGAGAGCCCGAGTCTTACTGGGCTGTCCAGT 4265
Db 155498 CTCGACTCTCCAGACAGGCTGTGTTTCTCCAGTACTCTGCGCGGAGGTCTTGA 155439
QY 4266 GCGAGGGAGTGGCTTGTGTGATCATTCGGGACAGTGGTCTGG 4314
Db 155438 CGCAGTGGGGGCGTCAATGTAGTGGACACATGGATACAGTTGTCTGG 155390

RESULT 5

US-09-105-537-34
; Sequence 34, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438051
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 4689
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-105-537-34

Query Match 0.9%; Score 40.8; DB 3; Length 4689;
Best Local Similarity 45.8%; Pred. No. 0.47;
Matches 141; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 3986 TCACCCACTGAGAGGAGGAGGAGCTGACCTCTTCTCCGAGAGCTGTTTC 4045
Db 2588 TCACCTTACTGCGGAGGCTGGGACAGGCTCCTACATGAGTGGGCGCCATCTCC 2647
QY 4046 ACTTCTGAGACCGCGGGAGTGGGAGATGCGCATTCGCACTCTGGAACAAATCTGGAAC 4105

Db 2648 CCAACCGCAACCGGCAACACCCCGAGTCCCACTTACGCTTTCAGACCGAGCGCTTCT 2707
QY 4106 CAGTGGGCGGCAAGACTTCTCCAACTTACCCACAGCGGAGGCAACTGAGTGGC 4165
Db 2708 GAGTGGAGAGCTTCCGCGCCACACAGCGCGGAGCACTGGGCTTACCGGCTGATGGA 2167
QY 4166 CCTTCTGAGAGCCCTTACTCTTACACCTTGGGAAAGCCGATTTCTCCAGACAGG 4225
Db 2768 AGCGCTGACGGGCTCGGCGCAGGCGGACCTGTCGGGCGGTGATGTCGCGGTGGGA 2827
QY 4226 CTGAAGAGCCCGGAGTGTCTTACTGGGCTGTCCAGTGGAGGAGTGGGCTTG 4285
Db 2828 GCGAGCCAGAGCCGAGTGTGGGCGCGCTGAGAGCCCGGAGCGAGTGCAGCTAC 2887
QY 4286 TGGTGGGC 4293
Db 2888 TGGAGCC 2895

RESULT 6

US-09-105-537-5
; Sequence 5, Application US/09105537A
; Patent No. 6265202
; GENERAL INFORMATION:
; APPLICANT: Sherman, D.H.
; APPLICANT: Liu, H.
; APPLICANT: Xue, Y.
; APPLICANT: Zhao, L.
; TITLE OF INVENTION: DNA encoding methymycin and pikromycin
; FILE REFERENCE: 600.438051
; CURRENT APPLICATION NUMBER: US/09/105,537A
; CURRENT FILING DATE: 1998-06-26
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 36778
; TYPE: DNA
; ORGANISM: Streptomyces venezuelae
US-09-105-537-5

Query Match 0.9%; Score 40.8; DB 3; Length 36778;
Best Local Similarity 45.8%; Pred. No. 1.6;
Matches 141; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 3986 TCACCCACTGAGAGGAGGAGGAGCTGACCTTCTTCTCCGAGAGCTGTTTC 4045
Db 29578 TCACCTTACTGCGGAGGCTGGGCGGAGGCTTCACTGAGTGGGCGCCATCTCC 29637
QY 4046 ACTTCTGAGACCGCGGAGTGGGAGATGCGCATTCGAGTGGAGCAATGCTGGAAC 4105
Db 29638 CACCGGCAACCGGCAACCGGAGCTTCCCACTTACGCTTTCAGACCGAGCGCTTCT 29697
QY 4106 CAGTGGGCGGCAAGACTTCTTCAACCACTTACCCAGAGCGGAGCAACTCAAGTGGC 4165
Db 29698 GAGTGGAGAGCTTCCGCGCCACACAGCGCGGAGAGTGGGTTACCGGTGAGTGGGA 29757
QY 4166 CCTTCTGAGAGCCCTTACTCTTACACCTTGGGAAAGCCGATTTCTCCAGACAGG 4225
Db 29758 AGCGCTGACGGGCTTCCGCGCAGGCGGAGCTTCCCGGCGGTGATGTCGCGGTGGGA 29817
QY 4226 CTGAAGAGCCCGGAGTGTCTTACTGGGCTGTCCAGTGGAGGAGTGGGCTTG 4285
Db 29818 GCGAGCCAGAGCCGAGTGTGGGCGCGCTGAGGCGCGGAGCGGAGGTGAGCTAC 29877
QY 4286 TGGTGGGC 4293
Db 29878 TGGAGCC 29885

RESULT 7

US-09-320-878-19
; Sequence 19, Application US/09320878A
; Patent No. 6117659

Query Match	0.9%;	Score 40.8;	DB 3;	Length 38506;
Best Local Similarity	45.8%;	Pred. No. 1.7;	Mismatches 167;	Indels 0;
Matches 141;	Conservative 0;	Mismatches 167;	Indels 0;	Gaps 0;
Qy 3986	TCACCCCATCTGAACGAGAGAGGGGACACTGACCTCACTCTTCTCCGAGACTGTTTC	4045		
Db 27720	TCACCTCATTGCGCGGAGCTTG66CCAAAGGCTTCACTGACCTGAGCGCCATCTCC	27772		
Qy 4046	ACTTCTAAGACCGGGGCGATGCGGAGATGGCCATTCGACTCTGGAAACAATCTGTGAAC	4105		
Db 27780	CCACCGGAACCGGCGACACACCCGAGGTCTCCCACTTACGCTTTCAGACGAGACGGCTTCT	27833		
Qy 4106	CAGTGGGCGCCGAAGACTACTCTCCAAACAATTCAACCCAGACCGACCAACTCAATGTC	4165		
Db 27840	GGCTGCAAGACTCCGCGCCCCACACAGCCGCCCCACACACTG6CTTACCGGTGAGTGA	27895		
Qy 4166	CCTCTCTCTGAAGGCTTACCTCTACACCTTGCGGAAACAGCCGATTGCTCCAGACGAG	4225		
Db 27900	AGCGCGTGAAGCGGCTCTCGGCGAAGCGGACCTGTCTCGGCGCGTGAATGTCGCGTGGGA	27953		
Qy 4226	CTGAAGAAGCCCCCGAGAGTGTCTACTAGGGCTGTCCAGTGGAGAGGGAGTGGGCTTG	4285		
Db 27960	GCGAGCCAGAAAGCCGAGAGTGTGGGCGGCTGAAGGCGCGGAGAGCGAGTGCAGCTAC	28015		
Qy 4286	TGATGGGC 4293			
Db 28020	TGGAAGCC 28027			
RESULT 8				
US-09-141-908-1				
; Sequence 1, Application US/09141908				
; Patent No. 6503741				
; GENERAL INFORMATION:				
; APPLICANT: ASHLEY, Gary				
; APPLICANT: BETLACH, Melanie C.				
; APPLICANT: BETLACH, Mary				
; APPLICANT: MCDANIEL, Robert				
; APPLICANT: TANG, Li				
; TITLE OF INVENTION: Combinatorial Polyketide Libraries Produced Using a				
; TITLE OF INVENTION: Modular PKS Gene Cluster as Scaffold				
; FILE REFERENCE: 300622002100				
; CURRENT APPLICATION NUMBER: US/09/141,908				

Query Match	0.9%	Score 40.8	DB 4	Length 38506
Best Local Similarity	45.8%	Pred. No. 1.7		
Matches 141	Conservative 0	Mismatches 167	Indels 0	Gaps 0
QY	3986	TCACCCCATGTAACGAGAGAGGGGACACTGACCTCACTCTTCTTCGAGACTGTTTC	4045	
DB	27720	TCACCTCACTCGCGCGAAGCCTGGGCCCAAGGCTCTCACTGCACTGGGGCCCATCTCC	27779	
QY	4046	ACCTTCAGACCGCGGGGCATGCGGATGGCCATGCGACCTCGGAACAATGTGTGAAC	4105	
DB	27780	CCACCCGACCGGACCAACACCCCGAAGCTCCCACTACGCTTCCAGACCGAGCGCTTC	27833	
QY	4106	CAGTGGGCGGACGACTACCTCCACAACTTCACACGACCGGACCAACTCAAGTGC	4165	
DB	27840	GGCTGCGAGACTCCGGGCCACACGACGCGCCGACGACACTGGGGTTACCCGCTGAGTGA	27899	
QY	4166	CTCTCTCTGAGAGCCCTTACCTTAACCTTGCGAACAACCGATGTCTCCAGACAG	4225	
DB	27900	AGCCGCTGACGCGCTTCGCGACAGCGGACCTGTCCGGGCGTGTGATGTCGCTGGGA	27955	
QY	4226	CTGAAGAAGCCCCGAGGTGCTTACTGAGCTGTCCAGTGGCAGCGGAGTGGGCTTG	4285	
DB	27960	GCGAGCGAAGAAGCCGAGCTCTGGGCGCGCTGAAGCCGCGGGAGGGAGTGAAGTAC	28015	
QY	4286	TGCTGGGC 4293		
DB	28020	TGGAAGCC 28027		
RESULT 9				
US-09-657-440-19				
Sequence 19, Application US/09657440				
Patent No. 6509455				
GENERAL INFORMATION:				
APPLICANT: ASHLEY, Gary				
APPLICANT: BETLACH, Melanie C.				
APPLICANT: BETLACH, Mary C.				
APPLICANT: MCDANIEL, Robert				
APPLICANT: YANG, Li				
TITLE OF INVENTION: RECOMBINANT NARBONOLIDE POLYKETIDE SYNTHASE				
FILE REFERENCE: 300622002120				
CURRENT APPLICATION NUMBER: US/09/657,440				
CURRENT FILING DATE: 2000-09-07				
PRIOR APPLICATION NUMBER: 09/320, 878				
PRIOR FILING DATE: 1999-05-27				
PRIOR APPLICATION NUMBER: CIP OF 09/141,908				
PRIOR FILING DATE: 1998-08-28				
NUMBER OF SEQ ID NOS: 34				
SOFTWARE: PatentIn Ver. 2.0				
SEQ ID NO 19				
LENGTH: 38506				
TYPE: DNA				
ORGANISM: Streptomyces venezuelae				
US-09-657-440-19				
Query Match 0.9%; Score 40.8; DB 4; Length 38506;				
Best Local Similarity 45.8%; Pred. No. 1.7;				
Matches 141; Conservative 0; Mismatches 167; Indels 0; Gaps 0;				

	Best Local Similarity	45.8%	Pred No. 1,7:		Mismatches	167;	Indels	0;	Gaps	0;
	Matches	141;	Conservative	0;						
Oy	3986	TCAACCCACTGAAAGAGAGAGGGGACACTGACCTCACTTCTTCGAGACTGTTC	4045							
Db	27720	TCACTCACTCGCGGAAGCCTGGGCCAAGGCCTCACATCATGAGTGAGGCCATCTCC	27779							
Oy	4046	ACTTTCTAGAACCAGGCGGCATGCCGAGTGGCCATGCGACTCTTGAAACCAATGCTGGAAC	4105							
Db	27780	CAACCCGAACCGGCGCACCAACCCGAGGTCTCCCACTACGCTTCCAGACCGAGCGCTTCT	27839							
Oy	4106	CAGTGGGCCCGAMAGTACTCTCAAACAATTCAACCAACGCGAGCCAAACTCAAGTGC	4165							
Db	27840	GGCTGCAGAGACTCCGCGCCACCAAGCGCGCCGACGACTGGCGTTACCGCGTGAATGA	27899							
Oy	4166	CCTTCTCTGAGAGCCCTTACCTCTACACCTCTGGGAAACAGCCGATTGCTCCAGACGAG	4225							
Db	27900	AGCGCGCTGAGCGGCTTCGCGCAAGCGGACCTGTCTCGGCGCGTAGATGTCTGCGCGGGA	27959							
Oy	4226	CTGAAAGAGCCCCGAGAGTCTCTACTGAGGCTGTCCAGTGGAGAGGGAGATCGGCTTG	4285							
Db	27960	GCGAGCCAAGAACCGAGCTGTGGGCGCGCTGAAGGCGCGGAGGCGGAGTGCAGCTAC	28019							
Oy	4286	TGCTGGGC	4293							
Db	28020	TGGAAGCC	28027							

```

RESULT 10
US-09-338-671-1
Sequence 1, Application US/09338671
Patent No. 6194638
GENERAL INFORMATION:
APPLICANT: Dhugga, Kanwarpal
APPLICANT: Fallis, Patricia Lynne
APPLICANT: Nichols, Scott E.
TITLE OF INVENTION: Alteration of Hemicellulose
TITLE OF INVENTION: Concentration in Plants
FILE REFERENCE: 0782
CURRENT APPLICATION NUMBER: US/09/338,671
CURRENT FILING DATE: 1999-06-22
EARLIER APPLICATION NUMBER: 60/090,416
EARLIER FILING DATE: 1998-06-23
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 1451
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: CDS
LOCATION: (100)...(1194)
US-09-338-671-1

Query Match          0.9%; Score 40.2; DB 3; Length 1465;
Best Local Similarity 54.4%; Pred. No. 0.34;
Matches 81; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 2905 GACTTCTCTGTGTGCTGTGACGCTCTTCTTCCAGAAATCCAGCTCCGTGCTGGCGGAT 2964
      |||||
DB 196 GACTTCTCTGTGAGATGTGGCGGCTCTTCTTCCAGGCTCTACACTCATCTGTCGCGAGAC 255
      |||||

QY 2965 GGGCTCCAGATAGTGTCCTTCTTTGGCCCGACGATCCAGCTCCAAATCAGAAATTCAC 3024
      |||||
DB 256 GGGCAGCCGACCCAGACCATCAAGGTGCCCCGAGGGCTTGCATCAGAACTTACAAACCG 315
      |||||

QY 3025 TCCGAGCTGGCCAGAGCCCTTTGGACCAA 3053
      |||||
DB 316 AACGACATCAACCGCATCTCTGGGCTCAA 344
      |||||

RESULT 11
PCT-US95-04801-4

```

```

; Sequence 4 Application PC/TUS9504801
; GENERAL INFORMATION:
; APPLICANT: Martin, Juan F.
; APPLICANT: Coque, Juan R.
; APPLICANT: Enguita, Francisco J.
; APPLICANT: Fuente, Juan L.
; APPLICANT: Llares, Francisco J.
; APPLICANT: Liras, Paloma
; TITLE OF INVENTION: DNA ENCODING CEPHARYNCIN BIOSYNTHESIS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John W. Wallen III
; STREET: P.O. Box 2000
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04801
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wallen III, John W.
; REGISTRATION NUMBER: 35,403
; REFERENCE/DOCKET NUMBER: 19179
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-4720
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 800 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; PCT-US95-04801-4

Query Match 0.9%; Score 39.6; DB 5; Length 800;
Best Local Similarity 50.8%; Pred. No. 0.36;
Matches 121; Conservative 0; Mismatches 114; Indels 3; Gaps 1;

QY 2691 CAACCTGCTGGAATTCTCTGAACCCCACTATCATCGGCGAGGTGTTCTCTGGAAACCCAGA 2750
    |||
DB 95 CACCCGCGAGGACTTCTCGGACCTC--AACCTTTCGGGGGGCTGGGGGGAGAGCCGGGT 151
    |||

QY 2751 CAAGTGGCCAGTGAGCAGGCGCAGGCGTGTGTGACTGCGTTCGACCCCTGGCGGAGAA 2810
    |||
DB 152 CTACACACCCGCGGTGCTGGCCGACCGCGCGGCGGACTGGCCGCTCGACCGGTGGCCGA 211
    |||

QY 2811 CTCCCAAGAGCTACCAAGGCTGAGGCGCTTCAGCGGAGCTTACCGGAGCAGACATGCGGA 2870
    |||
DB 212 GGCCCGCGGCGATCTCGGGTTCTCGCACTTGCCCGGCTACCAAGTGGGGGCGGCTGGCGAT 271
    |||

QY 2871 GCTGATGGGGCTAGGCGCGCTATGACACGACGAGAGGAGACTTCTCTGTGCTCTGACGCC 2928
    |||
DB 272 GCTGAAGAACCCGAGACACCCAGCGCGGCTTACAGACGACTGATGTGCAACTGCGGCGCC 329
    |||

RESULT 12
US-08-232-463-14/C
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Wallen III, John W.
; REGISTRATION NUMBER: 35,403
; REFERENCE/DOCKET NUMBER: 19179
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-3905
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 800 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
PCT-US95-04801-4

Query Match          0.9%; Score 39.6; DB 5; Length 800;
Best Local Similarity 50.8%; Pred. No. 0.36;
Matches 121; Conservative 0; Mismatches 114; Indels 3; Gaps 1;

QY      2691 CAACCTGTGACCTTCTCTGAACCCCACTATCATCTCGCAGGTGTTCTGTGAAACCCAGA 2750
      |||||
DB      95  CACCCGCAGGACCTTCTCGACCTC---AACCTTTCCGGGGGCTGGGGGAGACCCGGT 151
      |||||

QY      2751 CAATGCCCGATGAGAGAGCGACAGCGTGTGTAACGTGGTTCTGAACCCCTGGCGGAGAGA 2810
      |||||
DB      152 CTACACACCCCGCGGTGCTGGCCGACCGCGCGGAGACTGGCCGCTCCAGCCGCTGGGCCGA 211
      |||||

QY      2811 CTCCCAAGAGCTACCCAGAGCTGAGAGGCTTCAGCCGACCTACCCGAGACGATGCGGA 2870
      |||||
DB      212 GGCCCGCGCATCTCGGGTTCTCGAATTGCCCGCTACCAAGTGGCGGCGCTGGCGAT 271
      |||||

QY      2871 GCTGATGGGGTACAGGCCGCTATGACACGACGAGAGAACTTCTGTGTGCTCTGACGCC 2928
      |||||
DB      272 GCTGAAGAACCCGAGACCCAGCGCGCTACCAAGACCTGATGCTCGAATCGCGGCC 329
      |||||

RESULT 12
US-08-232-463-14/C
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS

```

NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base paire
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpc-Fls
US-08-232-463-14

Query Match 0.9%; Score 38; DB 1; Length 7218;
Best Local Similarity 4.5%; Pred. No. 3.8;
Matches 17; Conservative 197; Mismatches 162; Indels 0; Gaps 0;

QY 3479 GGGAGGGGACGAGCTAATGTGGACGAGGAGGCGCAGAGCTAGGACATGCGCAG 3558
DB 1410 RRR 1351
QY 3539 CCCAGGCTGGACCTGGTAGAGCAATGAATAAGCCCGACATCAACCTGGAGAAG 3598
DB 1350 RRR 1291
QY 3599 ACTGGAAGCTGGTCACTCTTCAATTGGGGTCAACGACTGTGTCAATTACTGTGAATC 3658
DB 1290 RRR 1231
QY 3659 CGGAGGCCACTTGGCCACGGAATATGTTGACGACATCCACAGCCCTGACATCTCT 3718
DB 1230 RRR 1171
QY 3719 CTGAGAGCTCCCAAGGCTTTCGTCACGCTGTGAGAGTATGAGCTGAGCTAGCCTGT 3778
DB 1170 RRR 1111
QY 3779 ACCAGGCGCAAGCGGGAATGTGCATGCTGCGAGCTCAGACAACATGCACTTGCTCA 3838
DB 1110 RRR 1051
QY 3839 GACACTGCGCAAGCTC 3854
DB 1050 ACTGACGCCAAGCTC 1035

RESULT 13
US-09-252-991A-5791
Sequence 5791, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 5791
LENGTH: 513
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5791

Query Match 0.9%; Score 37.6; DB 4; Length 513;
Best Local Similarity 49.5%; Pred. No. 1;
Matches 97; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 2790 CGTTTGACCTGGGGAGAACTCCCAAGAGCTAGCCAGGCTGGAGGCTTGACCGAGC 2849
DB 101 CGCCCTGTGGCTGGATCATCGCCCTGGCGGCTCGCTGTGCTCGGCTATGCCAC 160
QY 2850 CTACCGGAGAGCATGCGCGAGCTGTGGGGTCAAGCCGCTATGACAGCAGAGACTT 2909
DB 161 CGTCCGAAACCGCTGTATCGGGATCGCAACCGCTAGTGGAATCTTCGCAACGT 220
QY 2910 CTGTGTGTGTGAGCGCTTCTTCCAGAAATCAGCTCCCTGTCTCGGCGATGGGCT 2969
DB 221 GCCCTGTCTGTGAGCTGTTCATCTTCTCTGTACTCCGACCTGTGCCGAGGG 280
QY 2970 CCCAGATAGCTCTTC 2985
DB 281 CTGCGAGACTGTTC 296

RESULT 14
US-09-252-991A-5809
Sequence 5809, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 5809
LENGTH: 762
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-5809

Query Match 0.9%; Score 37.6; DB 4; Length 762;
Best Local Similarity 49.5%; Pred. No. 1.3;
Matches 97; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 2790 CGTTTGACCTGGGGAGAACTCCCAAGAGCTAGCCAGGCTTGACCGAGC 2849
DB 124 CGCCCTGTGGCTGGATCATCGCCCTGGCGGCTCGCTGTGCTCGGCTATGCCAC 183

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 6, 2004, 06:31:30 ; Search time 8433 Seconds
(without alignments)
12614.810 Million cell updates/sec

Title: US-10-054-691-1
Perfect score: 4377
Sequence: 1 atggggcctgcggcagcat.....tgcgcactgtggccctcctag 4377

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estln:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gse_hum:*
18: em_gse_inv:*
19: em_gse_pln:*
20: em_gse_vrt:*
21: em_gse_fun:*
22: em_gse_mam:*
23: em_gse_mus:*
24: em_gse_pro:*
25: em_gse_rtd:*
26: em_gse_rhg:*
27: em_gse_vrt1:*
28: gb_gse1:*
29: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1626	37.1	3148	11	AK030142 Mus muscu
2	868.8	19.8	935	9	AL542856 AL542856
3	788	18.0	1081	12	BM923066 BM923066
4	766.4	17.5	935	13	BU150503 BU150503

C 5	760.6	17.4	903	13	EX349938
C 6	732.2	16.7	974	13	EX349937
C 7	730.6	16.7	900	13	EX369782
C 8	712.6	16.3	861	13	BU183160
C 9	708.2	16.2	880	12	BI912651
C 10	687.8	15.7	1186	13	EX354650
C 11	677.4	15.5	1033	12	BM549919
C 12	671.4	15.3	763	12	BI912073
C 13	627.8	14.3	823	12	BM488437
C 14	607	13.9	607	12	BM714739
C 15	578	13.2	1562	11	AK015993
C 16	576.2	13.2	608	10	AW851123
C 17	467.8	10.7	587	9	AW752833
C 18	450.2	10.3	617	9	AI424952
C 19	419.8	9.6	978	13	BU540700
C 20	405	9.3	716	13	BU683311
C 21	401	9.2	625	14	CA447656
C 22	398.8	9.1	412	12	BM146939
C 23	372.2	8.5	583	9	AA418082
C 24	371.6	8.5	389	12	BM146940
C 25	365.4	8.3	616	9	AW467395
C 26	349.6	8.0	979	14	BU715569
C 27	344.4	7.9	570	10	BF510463
C 28	343	7.8	567	13	BU739708
C 29	339.4	7.8	575	13	BU743739
C 30	322.2	7.4	648	14	BY731604
C 31	322.2	7.4	2929	11	AK087596
C 32	317.4	7.3	541	10	BF854199
C 33	310.2	7.1	701	12	BI696117
C 34	303.8	6.9	667	10	BB616890
C 35	301.8	6.9	534	10	BE137831
C 36	301.4	6.9	525	10	BE684057
C 37	299.8	6.8	469	14	CB733804
C 38	299.8	6.8	513	4	BS17602
C 39	299.6	6.8	511	10	BE855594
C 40	294.4	6.7	511	10	BE645470
C 41	284.4	6.5	430	9	AA646625
C 42	282.2	6.4	495	9	AA461827
C 43	274.6	6.3	409	9	AA512437
C 44	263.4	6.0	725	13	BU633808
C 45	262.8	6.0	762	10	BF168176

ALIGNMENTS

RESULT 1
AK030142
LOCUS
DEFINITION Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4932703L18 product:weakly similar to PHOSPHOLIPASE B (Rattus norvegicus), full insert sequence.
ACCESSION AK030142
VERSION AK030142.1 GI:26081683
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carinci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
Carinci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL MEDLINE
2049374
PUBMED 11042159

REFERENCE
AUTHORS
3
Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kikunishi, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujimake, S., Inoue, K., Togawa, Y., Iwata, M., Ohara, E., Matsui, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--364-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
11076861

TITLE
JOURNAL
MEDLINE
PUBMED
20530913
11076861

REFERENCE
AUTHORS
4
Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Akakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamakawa, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadoya, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pease, G., Tomita, M.,
Quackenbush, J., Schrim, L.M., Staudt, F., Suzuki, R., Wagner, L.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baidarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bull, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, D., Mombauts, P., Nordone, P.,
Ringo, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seta, T., Shibata, Y., Storch, K.F., Suzuki, H.,
Toyooka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohzuki, S.,
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851

TITLE
JOURNAL
MEDLINE
PUBMED
11217851

REFERENCE
AUTHORS
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3148)

TITLE
JOURNAL
MEDLINE
PUBMED
11217851

COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken,
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>
Location/Qualifiers

FEATURES

source
1. .3148
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM:DB:4932703L18"
/db_xref="taxon:10090"
/clone="4932703L18"
/sex="male"
/issue_type="testis"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
misc_feature
1. .3148
/note="weakly similar to PHOSPHOLIPASE B [Rattus
norvegicus] (SPTRL054728, evidence: FASTA, 84%ID,
26.6%length, match=1158) "
BASE COUNT
856 a 819 c 756 g 717 t
ORIGIN
Query Match 37.1%; Score 1626; DB 11; Length 3148;
Best Local Similarity 75.7%; Pred. No. 0;
Matches 2074; Conservative 0; Mismatches 640; Indels 24; Gaps 4;
QY 52 CAAGGACCCCTGAGATCATCTCTCTGAGAAAGATGATGAGAGGCGACTATGG 111
82 CAAGGGCCCTCCAAATCCATGCTCTTCTGAGAGAACATTTGGCATGCTCCAG 141
QY 112 CCAGA-----GACCTGGAAGATTTCTCCATGCCATGCAACCCAAATTAATAGAGTG 165
142 CAAGTGTTTTGACCCCTGAAGAAATTTCCATTTCTTGCAGGCAAGCAAGATTAAGATG 201
QY 166 AATATGCTTCTTAATACAGTCTCACTCTCTGAAAGCTTTGATTAATTTGGGAGCC 225
202 AGTGTCTTTCTGAGTCACTCTCTCTGAGAGCTCTGAGACATTAATCTGCGAGCC 261
QY 226 ATTGCAATCTGAATTTCTCTCCAGACCGAGGAGCGGCGATCTGAGAG----- 276
262 ATCGGCAACCGAAATTTCTCTAGCCCTGGGCTGGCAGATCAACATGAGAAACT 321
QY 277 CAAGACTGAGCTGAAGGCGCAGAGCTGTGATGAGATGATGAGATGATGAGATGAGATG 336
322 CAAGCAATCAAGAACCAACCAAGAGATGATGATGAGATGATGAGATGATGAGATG 361
QY 337 ATCATCATATTTTCACT 396
382 ATCATCATATTTTCACT 441
QY 397 ATACCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 456
442 GCACTCACTACTCTCTGAAAGATTTGTGATTAAGGCTGAGAGGCGCTG 501
QY 457 AAGAGAACTGCACTTGAATTTGAATTTGAATTTGAATTTGAATTTGAATTTGAATTTGA 516
502 AAGGACAAACCGGAACTGATTTGAAAGAGATGAAATCTCACTGTTCTTCACT 561
QY 517 AATGCAAGCAGTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 576
562 AATGCAAGCAGTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 621
QY 577 GTGATGAGCTGATGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 636
622 ATGAGATGCTGAGGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 681
QY 637 AACCTGATGCTCTCTGAGGTTGAGAGGCTCTCTGATGATGATGATGATGATGATGATG 696
682 AATTGTGATGCTCTCTGAGGTTGAGAGGCTCTCTGATGATGATGATGATGATGATGATG 741
QY 697 AGCCGCAACAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 756
742 AGCCGCAACAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 798
QY 757 ATGCAATGCTTATGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTG 816
799 ATGCAATGCTTATGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTG 858
DB

QY 817 CAGAGTCTTCAACCGTGGTTTTCAGACCTTCTTATGAGCAACCCCATCTTAC 876
 DB 859 CATTGACCTTGGCGGTGGTTTTCAGACCTTCTTATGAGCAACCCCATCTTAC 912
 QY 877 TGGAGAGACCCCGACTCCAGATTTTCAACGCTGGCTGGCATCTTGAATAGATG 936
 DB 913 TTGAAAAGGTCTCGCCCCAGATGCCACCACTGGCTTCAAGATCTTGAATAGATG 972
 QY 937 ATGGACCCAGAGAGAGAAAGATGAGCCATTGAGTGTAAACAAGGAGCCATGAA 996
 DB 973 ATGAAACCGAGGTGCTAAAGAGATGAGCTAATAATCTGAGAGAGAAAGAAATGAA 1032
 QY 997 TGTCCCTCAGAGAGACCCCTATCTGTTCAGTTCAGAAACAGCACTCTGACAGA 1056
 DB 1033 TGTCCCTCAGAGAGAGTCCCTATCTGTTCAGTTCAGAAATGAGCACTTCCAGGCCAGA 1092
 QY 1057 CTGCAAGAACCCCAAGACAACTTGAAGTGAAGAGAGAGCCGAAATCAGATGCTGAC 1116
 DB 1093 CGACTGAAACCTTAAACAAAGCTTCAGATGAAAGAGATCAGAAATTCACCTGCTGAC 1152
 QY 1117 AAAGACCCCTCCGATAGGTTCACCTCAGTTCATAGGTGAAGCCGGCTGACATCAAC 1176
 DB 1153 AAGAACCCCTCGAACTCAATCCCAACAGCTTCAGAGCTTGGCGCAGCTGATATCAAG 1212
 QY 1177 GTAATTGAGACCCCTGGGTGACTCTCTCAAGGAGCAATGGGGCCGGGTCCACACCTGGG 1236
 DB 1213 ATCATAGAGACATTTGGGTGACTGCTCAAGGAGCAATGAGCTGGGGCCAGCCCTTGG 1272
 QY 1237 AAGCTTTGAGACCTTGAAGTCACTGATACGAGGCTGTCTGAGAGCTGGCGAGATGAG 1296
 DB 1273 AATATCTTGGATGTCTTAAGTGAATCCAGAGCTGTCTGAGATGGCGGGATGAG 1332
 QY 1297 AAGATGGGACCTGTACACCTCTGGGAGACATCTCTGGGAAATTCACCTTCCCTGAG 1356
 DB 1333 ACCATGAAACCGGTGACACCTTCAACCAATCTCTGGGAAATTCACCTTCCCTGAG 1392
 QY 1357 GGGCTTCTGTGGGACTGGGAAAGAAACAGTCTTAATGCTTCTTAAACAGGCGTG 1416
 DB 1393 GGGCTTCTGTGGGACTGGGAAAGAAAGCACTTCCAGACATCTTCAACAGGCGTG 1452
 QY 1417 GCAGAGGCGGAGCTGAGATCTAAGCTGTCCAGGAGGAGGCTGGTGAAGCTGATGAG 1476
 DB 1453 GCAGAGGCGCAATCTGATGGCTTAACTGCGCAGGCGCAAGAGCTGGTGAAGCTGATGAG 1512
 QY 1477 AATGACAGAGATCACTTTCAGAAAGCTGGAAGATTAATCCTGTTATAGCGGCG 1536
 DB 1513 GCAGAGAGCAATTAATCTTTCAGAAAGCTGGAAGATTAATCCTGTTATAGAGGCG 1572
 QY 1537 AATGACCTCTGATTTCTGCAATGATCTGCTCAGATTTCTCCAGAACTTCAAGAC 1596
 DB 1573 AATGACCTCTGATTTCTGCAATGATCTGCTCAGATTTCTCCAGAACTTCAAGAC 1632
 QY 1597 AACATTGAAAGGCGCTGAGATCTCTCAATGCTGAGAGTTCCTGGGCAATTTGAACTG 1656
 DB 1633 AACATCAAGAGCGCTGAGATCTCTCAATGCTGAGAGTTCCTGGGCAATTTGAACTG 1692
 QY 1657 GTGACGCTGCTGAGATCTCTCAATGCTGAGAGCTGTAACAGAGAAAGTCTAAGT 1716
 DB 1693 GCAATGCTGATGAGATCAACCTTTCAGAAAGCTTCAATGAACTTCAAGCTCATG 1752
 QY 1717 CCAAGATGATCTCTGAGTCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1776
 DB 1753 CCAAGATGATCTCTGAGTCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1812
 QY 1777 GAATCTGCTAAGCTTCAAGATTTCAAGAAATTTCAAGAGAGAGAGAGAGAGAGAGAG 1836
 DB 1813 GAATCTTCAAGCTTCTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1872
 QY 1837 GAGAGTGGCGATATGACACAGAGAGAGATTTTACTGTGTGTGAGAGCGGTTCTTGA 1896
 DB 1873 AATAGCGGAGAGATATGACACAGAGAGATTTTACTGTGTGTGAGAGCGGTTCTTGA 1932

QY 1897 AACGTGAGATGCGCAAGACCTTCGAGAGATTCCTGACAACTTCTTGGCTCTGAC 1956
 DB 1933 AATGTTCATGCGACAGAGACCCCGAGAGGCGTCCGAGCAAGTCTTCTTGGCCCTGAC 1992
 QY 1957 TGTTCCTTCAAGAGAGAGTCTTCACTCCGAGAGCCGAGTCTCTTGAAGCAATATG 2016
 DB 1993 TGTTCCTTCAAGAGAGAGTCTTCACTCCGAGAGCCGAGTCTCTTGAAGCAATATG 2052
 QY 2017 CTGAGACCTTGGGCGAGAGAGAGTCTTCACTTGAAGAGAGAGAGAGAGAGAGAGAG 2076
 DB 2053 CTGAGACCTTGGGCGAGAGAGAGTCTTCACTTGAAGAGAGAGAGAGAGAGAGAGAG 2112
 QY 2077 TGTCCGAGAGAGAGAGAGAGTCTTCACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 2136
 DB 2113 TGTCCGAGAGAGAGAGAGAGTCTTCACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 2172
 QY 2137 ACTGGCTGATGAGAGAGAGAGAGTCTTCACTTGAAGAGAGAGAGAGAGAGAGAGAG 2196
 DB 2173 ACTGGCTGATGAGAGAGAGAGAGTCTTCACTTGAAGAGAGAGAGAGAGAGAGAGAG 2232
 QY 2197 CTGAGACCTGAGAGAGAGAGAGAGTCTTCACTTGAAGAGAGAGAGAGAGAGAGAGAG 2256
 DB 2233 CTGAGACCTGAGAGAGAGAGAGAGTCTTCACTTGAAGAGAGAGAGAGAGAGAGAGAG 2292
 QY 2257 GGAATGAGCTCCAGAGAGAGAGAGAGTCTTCACTTGAAGAGAGAGAGAGAGAGAGAGAG 2316
 DB 2293 GGAATGAGCTCCAGAGAGAGAGAGAGTCTTCACTTGAAGAGAGAGAGAGAGAGAGAGAG 2352
 QY 2317 TACAGTGAAGAGAGAGAGAGAGTCTTCACTTGAAGAGAGAGAGAGAGAGAGAGAGAG 2376
 DB 2353 TATAGTGTGAGAGAGAGAGAGAGTCTTCACTTGAAGAGAGAGAGAGAGAGAGAGAGAG 2412
 QY 2377 GAGTTTAAAGAAAGTCTCAAGAGTCTTCACTTGAAGAGAGAGAGAGAGAGAGAGAGAG 2436
 DB 2413 AATTTTAAAGAAAGTCTCAAGAGTCTTCACTTGAAGAGAGAGAGAGAGAGAGAGAGAG 2472
 QY 2437 GCATTCCTCAATCAAGAGTCTTCCGAGAGAGAGAGAGTCTTGAAGAGAGAGAGAGAG 2496
 DB 2473 GCTTTCCTTCAAGAGTCTTCCGAGAGAGAGAGAGTCTTGAAGAGAGAGAGAGAGAGAG 2532
 QY 2497 ACTTGATGAGAGAGAGAGAGAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2556
 DB 2533 ACTTGATGAGAGAGAGAGAGAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2592
 QY 2557 ATCAAGTCTGATGAGAGAGAGAGAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2616
 DB 2593 ATCAAGTCTGATGAGAGAGAGAGAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2652
 QY 2617 TGTGAGAGAGAGAGAGAGAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2676
 DB 2653 TGTGAGAGAGAGAGAGAGAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2712
 QY 2677 CCCAGAGTCTGATGAGAGAGAGAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2736
 DB 2713 CCCAGAGTCTGATGAGAGAGAGAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2772
 QY 2737 CTGAGAGAGAGAGAGAGAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2774
 DB 2773 CTGAGAGAGAGAGAGAGAGTCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2810
 RESULT 2
 AL542856 935 bp mRNA linear EST 12-MAY-2003
 LOCUS AL542856 Homo sapiens PLACENTA Homo sapiens cDNA clone CSDB0133YE22
 DEFINITION 5-PRIME, mRNA sequence.
 ACCESSION AL542856
 VERSION AL542856.2 GI:30548429
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 935)
 AUTHORS Li, M. B., Gruber, C., Jesse, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT On Feb 15, 2001 this sequence version replaced gi:12875311.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 1474.f
 Contact: Feng Liang Email: fliang@life.techn.com URL:
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID: CS00E013B011091.
 FEATURES
 source
 1..935
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS00E013YE23"
 /issue_type="PLACENTA"
 /clone_id="Homo sapiens PLACENTA"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo (dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."
 BASE COUNT 229 a 272 c 262 g 172 t
 ORIGIN
 Query Match 19.8%; Score 868.8; DB 9; Length 935;
 Best Local Similarity 98.1%; Pred. No. 7.5e-215;
 Matches 879; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
 QY 3477 CTGGAGAGGAGACAGCAGACTAAATGTGGCAGCGAAGGGGCCAGAGCTAGAGACATGCC 3536
 DB 40 CTGGTACCGGTCCGGAATTCCTCCGGATGGCGAGAGGGGCCAGAGCTAGAGACATGCC 99
 QY 3537 AGCCAGAGGCTGGGACCTGTGTAGAGCGAATGAAAACGCCGCCCATCACTGGAGAA 3596
 DB 100 AGCCAGAGGCTGGGACCTGTGTAGAGCGAATGAAAACGCCGCCCATCACTGGAGAA 159
 QY 3597 AGACTGAGAGCTGTACACTCTTCATTTGGGGTCAACGACTTGTGTCATTACTGTGAAA 3656
 DB 160 AGACTGAGAGCTGTACACTCTTCATTTGGGGTCAACGACTTGTGTCATTACTGTGAAA 219
 QY 3657 TCCGAGAGGCCACTTGGCCAGCAATATGTTCAGACATCCACAGGCCCTGACATCCT 3716
 DB 220 TCCGAGAGGCCACTTGGCCAGCAATATGTTCAGACATCCACAGGCCCTGACATCCT 279
 QY 3717 CTCTGAGAGAGTCCCAAGGGTTTCGTCAAGTGTGTGAGGTCTATGAGCTGGTACGCT 3776
 DB 280 CTCTGAGAGAGTCCCAAGGGTTTCGTCAAGTGTGTGAGGTCTATGAGCTGGTACGCT 339
 QY 3777 GTACACAGGCGCAAGGCGGAAATGTGCATCTGGAGCTCAGAACTGCACTTGCT 3836
 DB 340 GTACACAGGCGCAAGGCGGAAATGTGCATCTGGAGCTCAGAACTGCACTTGCT 399
 QY 3837 CAGACACTCGCAAGGCTCCCTGAGAGCAAGAACTGAAGAACTGAACCTTCCA 3896
 DB 400 CAGACACTCGCAAGGCTCCCTGAGAGCAAGAACTGAAGAACTGAACCTTCCA 459
 QY 3897 GCATGGCATCTCCAGTTTCTCTACTGGCAACATATACAGAGGTGAGACTTTGGCT 3956
 DB 460 GCATGGCATCTCCAGTTTCTCTACTGGCAACATATACAGAGGTGAGACTTTGGCT 519
 QY 3957 TGTGTGAGAGGCTTCTTCCAAAACACATCACCCTGTAAGAGAGAGGAGCACTGA 4016
 DB 520 TGTGTGAGAGGCTTCTTCCAAAACACATCACCCTGTAAGAGAGAGGAGCACTGA 579
 QY 4017 CTTGACCTTCTTCTCGAGAGACTGTTTCACTTCTCAGACCGCGGCAATGCCAGATGCC 4076
 DB 580 CTTGACCTTCTTCTCGAGAGACTGTTTCACTTCTCAGACCGCGGCAATGCCAGATGCC 639

QY 4077 CATGCACTCTGGAAACAACATGCTGGAACAGTGGGCCGCAAGACTTCCAACTT 4136
 DB 640 CATGCACTCTGGAAACAACATGCTGGAACAGTGGGCCGCAAGACTTCCAACTT 699
 QY 4137 CACCACAGCCGAGCCAACTCAAGTGCCTCTCTGAGAGCCCTTACTTACACCT 4196
 DB 700 CACCACAGCCGAGCCAACTCAAGTGCCTCTCTGAGAGCCCTTACTTACACCT 759
 QY 4197 GCGGAACAGCCGAGTGTCTCCCAACAGGTGAAAGAACCCCGAGGTGCTTACTGGGC 4256
 DB 760 GCGGAACAGCCGAGTGTCTCCCAACAGGTGAAAGAACCCCGAGGTGCTTACTGGGC 819
 QY 4257 TGTCCAGTGGCAGCCGAGTGTGAGTGTGAGGAGCATATGAGGAGAGTGTGAG 4316
 DB 820 TGTCCAGTGGCAGCCGAGTGTGAGTGTGAGGAGCATATGAGGAGAGTGTGAG 879
 QY 4317 GTCCAGAGAGGTGGCCGAGGAGAGATTCCTCAATAGCCTGCGCATGTGCCC 4372
 DB 880 GTCCAGAGAGGTGGCCGAGGAGAGATTCCTCAATAGCCTGCGCATGTGCCC 935
 RESULT 3
 BM923066
 LOCUS
 DEFINITION
 5' mRNA sequence.
 ACCESSION
 BM923066
 VERSION
 BM923066.1 GI:19373445
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 1081)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LMNL at:
 http://image.llnl.gov
 Plate: LHM42798 row: j column: 10
 High quality sequence stop: 664.
 FEATURES
 source
 1..1081
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5757033"
 /issue_type="leukocyte"
 /lab_host="DH10B"
 /clone_id="NIH_MGC_118"
 /note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV
 (destroyed); RNA source leukocytes from anonymous pool of
 non-activated adult donors. Library is oligo-dT primed
 and directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 1.2-3.3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 027. Note:
 this is a NIH MGC Library."
 BASE COUNT 255 a 329 c 286 g 208 t 3 others
 ORIGIN
 Query Match 18.0%; Score 788; DB 12; Length 1081;
 Best Local Similarity 93.1%; Pred. No. 9.7e-194;
 Matches 925; Conservative 0; Mismatches 52; Indels 17; Gaps 9;

```

QY 2648 ATGCTTGAAGCTCTTCATAGAGAGGTGCGCAGAGTCTGTGCAACCTGTGAGACTTCC 2707
Db 16 ATGCTTGAAGCTCTTCATAGAGAGGTGCGCAGAGTCTGTGCAACCTGTGAGACTTCC 75
QY 2708 TGAACCCCACTATCATGCGGAGGTGTTCTCTGAGAAACCAAGACAGTCCCACTGAGC 2767
Db 76 TGAACCCCACTATCATGCGGAGGTGTTCTCTGAGAAACCAAGACAGTCCCACTGAGC 135
QY 2768 AGGCGACGCTTGTGTATGCTGCTGACCCCTGCGGAGAACTCCCAAGACTTACCA 2827
Db 136 AGGCGACGCTTGTGTATGCTGCTGACCCCTGCGGAGAACTCCCAAGACTTACCA 195
QY 2828 GGTGTGAGGCTTTCAGCCGAGCTTACCGGAGCAGACATGCGGAGCTGTGTGAGGCTCAGGCC 2887
Db 196 GGTGTGAGGCTTTCAGCCGAGCTTACCGGAGCAGACATGCGGAGCTGTGTGAGGCTCAGGCC 255
QY 2888 GCTATGACAGCAGAGAGGACTTCTGTGTGCTGACGCCCTTCTTCAGAACTCCAGC 2947
Db 256 GCTATGACAGCAGAGAGGACTTCTGTGTGCTGACGCCCTTCTTCAGAACTCCAGC 315
QY 2948 TCCCTGTCTGCGGAGATGGGCTCCCAATAGCTCTTCTTGGCCAGACTGATCCACC 3007
Db 316 TCCCTGTCTGCGGAGATGGGCTCCCAATAGCTCTTCTTGGCCAGACTGATCCACC 375
QY 3008 CAATATGAAATTCCTATCCAGCTGCGCAGAGCCCTTGGACCAATATGCTTGAACAC 3067
Db 376 CAATATGAAATTCCTATCCAGCTGCGCAGAGCCCTTGGACCAATATGCTTGAACAC 435
QY 3068 TTGGAACCAAAACAGAGACCTGTGACCTGAGAGAGAGATGCCCATGCTGTCCATC 3127
Db 436 TTGGAACCAAAACAGAGACCTGTGACCTGAGAGAGAGATGCCCATGCTGTCCATC 495
QY 3128 AGAATGAGCCCTTCTCTGAAACCCCTGTGAAATGATTACTACAGTACCCATCAAGCAG 3187
Db 496 AGAATGAGCCCTTCTCTGAAACCCCTGTGAAATGATTACTACAGTACCCATCAAGCAG 555
QY 3188 CCAATGGAAGCTGGGAGTGAATCTCTGTGTACAGAGTGAAGGCTTCCAAATGATGTTG 3247
Db 556 CCAATGGAAGCTGGGAGTGAATCTCTGTGTACAGAGTGAAGGCTTCCAAATGATGTTG 615
QY 3248 CAACTCTGTGCAACAGCTCCGACAGCAGACATCAAAAGTGTGCGG - CCTGGGTGAC 3306
Db 616 CAACTCTGTGCAACAGCTCCGACAGCAGACATCAAAAGTGTGCGG - CCTGGGTGAC 675
QY 3307 TCTCTGACTACAGAGTGGAGCTGACCAAACTCTCAGTGAATCTTCACTTCTTGG 3366
Db 676 TCTCTGACTACAGAGTGGAGCTGACCAAACTCTCAGTGAATCTTCACTTCTTGG 735
QY 3367 AGGGAATCTCTTGGAGCACTTGGAGGAGTGAAGTGGAGTCACTAC - ACATGTC 3424
Db 736 AGGGAATCTCTTGGAGCACTTGGAGGAGTGAAGTGGAGTCACTAC - ACATGTC 795
QY 3425 CCAACATTTGAAAGAACTCAACCTT - ACCTCTTGGCTTCTCTACAGACCTGGG - 3481
Db 796 CCAACATTTGAAAGAACTCAACCTTCTTACTCTTCTTGGGCTTCTCTACAGACCTGGG 855
QY 3482 AGGGAAGCAGAG - ACTTAAATGTGG - - CAGCGGAAGGGGCGAGAGTGAAGG - ACATGCC 3536
Db 856 AGGGAAGCAGAG - ACTTAAATGTGG - - CAGCGGAAGGGGCGAGAGTGAAGG - ACATGCC 915
QY 3537 AGGCCAGGCC - - - TGGACCTGTGTAGAGCGAATGAAAA - - - ACAGCCCCCAGATCACT 3590
Db 916 AGGCCAGGCC - - - TGGACCTGTGTAGAGCGAATGAAAA - - - ACAGCCCCCAGATCACT 975
QY 3591 GGAGAAAGACTGGAAGCTGTGCACTTCTTAT 3624
Db 976 GGAGAAAGACTGGAAGCTGTGCACTTCTTAT 1009

```

```

RESULT 4
BU150503
LOCUS BU150503 935 bp mRNA linear EST 03-SEP-2002

```

```

DEFINITION AGENCOURT_8125945_Lupeki_dorsal_root_ganglion_Homo_sapiens_cDNA
ACCESSION BU150503
VERSION BU150503.1 GI:22664035
KEYWORDS EST
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 935)
AUTHORS NIH-MGC. http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strauberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LNL1355 row: n column: 23
High quality sequence start: 53
High quality sequence stop: 532.
Location/Qualifiers
1. 935
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:617622"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/clone_1b="Lupski_dorsal root ganglion"
/notes="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
Note1: Site 2: Sali; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCAGCCGCTCCG-3' and
5'-GACTGATTTAGATCGGAGCGGCGCCCTT(15)-3'. Size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
BASE COUNT 235 a 290 c 248 g 162 t
ORIGIN
Query Match 17.5%; Score 766.4; DB 13; Length 935;
Best Local Similarity 95.0%; P-Val 3.7e-188;
Matches 857; Conservative 0; Mismatches 36; Indels 9; Gaps 6;
QY 3484 GGGACAGCAGAGTAATGTGCGAGCGAAGGGGCGAGCTGAGGACATGCCAGCCAG 3543
Db 1 GGGACAGCAGAGTAATGTGCGAGCGAAGGGGCGAGCTGAGGACATGCCAGCCAG 60
QY 3544 GCCTGGAGCTGTGAGAGCGAATGAAAAAGCCCGACATCACTGAGAAAGACTGG 3603
Db 61 GCCTGGAGCTGTGAGAGCGAATGAAAAAGCCCGACATCACTGAGAAAGACTGG 120
QY 3604 AACCTGTCACTCTTCACTTGGGCTCAAGACTGTGTATTAATCTGAAATCCGAG 3663
Db 121 AACCTGTCACTCTTCACTTGGGCTCAAGACTGTGTATTAATCTGAAATCCGAG 180
QY 3664 GCCCACTGGCCAGGAATATGTTACAGACATCCAAGGCGCTGAGCATCTCTGAG 3723
Db 181 GCCCACTGGCCAGGAATATGTTACAGACATCCAAGGCGCTGAGCATCTCTGAG 240
QY 3724 GAGCTCCCAAGGGCTTTCGTCACGTTGTGAGGCTCATGAGCTGCTGATACAG 3783
Db 241 GAGCTCCCAAGGGCTTTCGTCACGTTGTGAGGCTCATGAGCTGCTGATACAG 300

```

QY 3784 GGCCAAAGCGGGAATATGTCATCTGAGCTGAGCACTGCTCAGACAC 3843
 Db 301 GGCCAAAGCGGGAATATGTCATCTGAGCTGAGCACTGCTCAGACAC 360
 QY 3844 TCGCAAGCTCTCTGAGAGCAAGAACTGAAAGTGAACCTTCACATGCG 3903
 Db 361 TCGCAAGCTCTCTGAGAGCAAGAACTGAAAGTGAACCTTCACATGCG 420
 QY 3904 ATCTCCAGTTTCTCTCTGAGCACTGAGCACTGAGGAGCTTGGGTTGAGTG 3963
 Db 421 ATCTCCAGTTTCTCTCTGAGCACTGAGCACTGAGGAGCTTGGGTTGAGTG 480
 QY 3964 CAGCCTTCTTCCAAACACACTGACCCCACTGAGAG-AGAGGGGACACTGACTCAC 4022
 Db 481 CAGCCTTCTTCCAAACACACTGACCCCACTGAGAGAGAGGGGACACTGACTCAC 540
 QY 4023 CTCTCTCTCGAGAGAGCTGTTTCACTTCTGAGACCGGCGGAGTCCGAGATGCGCATGCG 4082
 Db 541 CTCTCTCTCGAGAGAGCTGTTTCACTTCTGAGACCGGCGGAGTCCGAGATGCGCATGCG 600
 QY 4083 ACTCTGAAACAACATGCTGGAACCAAGTGGGCGGCAACATCTTCCAACTTCAACCA 4142
 Db 601 ACTCTGAAACAACATGCTGGAACCAAGTGGGCGGCAACATCTTCCAACTTCAACCA 660
 QY 4143 CAGCGGAGCCAAACTCAAGTGCCTCTCTGAGAGGCTTAACTTCAACCTGCGG-A 4201
 Db 661 CAGCGGAGCCAAACTCAAGTGCCTCTCTGAGAGGCTTAACTTCAACCTGCGGAA 720
 QY 4202 ACAAGCATTGCTTCCAGACAGGCTGAAAGAGCCCGAGAGTCTTACTGAGGCTGTC 4261
 Db 721 ACAAGCATTGCTTCCAGACAGGCTGAAAGAGCCCGAGAGTCTTACTGAGGCTGTC 780
 QY 4262 CAGTGGGA--GCGGAGTCCGCTTGTGGGGAGATCATCGGCACT-GGTCTGAGGT 4318
 Db 781 CAGGAGGAGGAGGAGGAGTCCGCTTGTGGGGAGATCATCGGCACTGGGAGGT 840
 QY 4319 GCAGAGAGGAGTGG---CCGAGAGGAAATCTTCCAA-TGAGCCCTGCGCACTGAGGCTC 4374
 Db 841 GCAGAGAGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
 QY 4375 TA 4376
 Db 901 CA 902
 RESULT 5
 BX349938/c 903 bp mRNA linear EST 05-MAY-2003
 LOCUS BX349938 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 DEFINITION cDNA clone CS0DC023YA08 3-PRIME, mRNA sequence.
 ACCESSION BX349938
 VERSION BX349938.1 GI:30371420
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 903)
 AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 1474.f,
 Contact: Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0BA10052A12_CS00473_2.
 Location/Qualifiers
 1..903
 /organism="Homo sapiens"

/mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DC023YA08"
 /issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
 /clone_id="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-0.190 (dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 162 a 242 c 269 g 223 t 7 others
 ORIGIN
 Query Match 17.4%; Score 760.6; DB 13; Length 903;
 Best Local Similarity 92.6%; Pred. No. 1,2e-186;
 Matches 833; Conservative 0; Mismatches 26; Indels 41; Gaps 2;
 QY 3431 TTCTGAAGAAATTCAACCTTACTCTCTTGGCTTCTCTACAGACACTGGAGGGGACAG 3490
 Db 903 TTGGAAGAAAGTTCAACCTTACTCTCTTGGCTTCTCTACAGACACTGGAGGGGACAG 844
 QY 3491 CAGGACTAAATGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3550
 Db 843 CAGGACTAAATGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 784
 QY 3551 ACCTGGTAGAGCAATGAAAAACAGCCCGGACATCAACTGAGAGAAAGTGGAGACTGG 3610
 Db 783 CCTGTGTAGAGCAATGAAAAACAGCNCAGACATCAACTGAGAGAAAGTGGAGACTGG 724
 QY 3611 TCAACTCTTCAATTGGGGTCAACGACTGTGTCTTATTACTGTGAAATCC----- 3659
 Db 723 TCAACTCTTCAATTGGGGTCAACGACTGTGTCTTATTACTGTGAAATCCGGTAGGCCCCC 664
 QY 3660 -----GAGAGCCCACTTGGCCAGGAAATATGTTTCAG 3690
 Db 663 GACCAACCCCATGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 604
 QY 3691 CACATCCAAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3750
 Db 603 CACATCCAAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 544
 QY 3751 GTGAGGTCATGAGAGTGGGCTGATGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3810
 Db 543 GTGAGGTCATGAGAGTGGGCTGATGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 484
 QY 3811 GCAGCTCAGAAACAATGCACTTGGCTCAGACATCTCCCAAAAGCTCCCTGGAGAGCAAGAA 3870
 Db 483 GCAGCTCAGAAACAATGCACTTGGCTCAGACATCTCCCAAAAGCTCCCTGGAGAGCAAGAA 424
 QY 3871 CTGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 3930
 Db 423 CTGAAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 364
 QY 3931 TACACAAGGCTGAGGAGCTTGGGTTGTGTGAGAGCTTTTCTTCCAAAACACACTGACC 3990
 Db 363 TACACAAGGCTGAGGAGCTTGGGTTGTGTGAGAGCTTTTCTTCCAAAACACACTGACC 304
 QY 3991 CCACTGAAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4050
 Db 303 CCACTGAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 244
 QY 4051 TCAGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4110
 Db 243 TCAGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 184
 QY 4111 GGCCGGAAGACTACTTCAACAACTTACCCAGGCGAGGCAAACTCAAGTGGCCCTCT 4170
 Db 183 GGCCGGAAGACTACTTCAACAACTTACCCAGGCGAGGCAAACTCAAGTGGCCCTCT 124
 QY 4171 CTTGAAGGCTTACTCTTCAACCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4230
 Db 123 CTTGAAGGCTTACTCTTCAACCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 64
 QY 4231 GAAGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4290

/note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

BASE COUNT 225 a 264 c 235 g 165 t 11 others

Query Match 16.7%; Score 730.6; DB 13; Length 900;
Best Local Similarity 91.1%; Pred. No. 7,8e-179;
Matches 821; Conservative 0; Mismatches 34; Indels 46; Gaps 3;

3324 GGGAGCTGCAACCAACTCTCACTGACCTTACCACTCTTGGAGGGGACTCTTGGAG 3383
5 GGGAGCTGCGCCCAACACTCTCCGTCGACCTTACCACTCTTGGAGGGGACTCTTGGAG 64
3384 CATTGGAGGGGAGTGGGAATCTTGGAGCTCACACACACTGCGCCCAACTCTTGGAGGAGT 3443
65 CATTGGAGGGGAGTGGGAATCTTGGAGCTCACACACACTGCGCCCAACTCTTGGAGGAGT 124
3444 CAACCTTACCTCTTGGAGCTTCTTACACAGCACTTGGAGGGGAGGAGGAGGAGGAGT 3503
125 CAACCTTACCTCTTGGAGCTTCTTACACAGCACTTGGAGGGGAGGAGGAGGAGGAGT 184
3504 GGCAGCGGAGGGGCGCAGAGCTAGGAGCATGCGAGCCGAGGCTGGAGGAGGAGGAGGAG 3563
185 GGCAGCGGAGGGGCGCAGAGCTAGGAGCATGCGAGCCGAGGCTGGAGGAGGAGGAGGAG 244
3564 AATGAAAAAAGAGGCGGAGCATCACTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 3623
245 AATGAAAAAAGAGGCGGAGCATCACTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 304
3624 TGGGGTCAAGCACTTGTGTCTTCTTGTGAGTCTTGTGAGTCTTGTGAGTCTTGTGAGT 3659
305 TGGGGTCAAGCACTTGTGTCTTCTTGTGAGTCTTGTGAGTCTTGTGAGTCTTGTGAGT 364
3660 -----GAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3703
365 GGCAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 424
3704 CCTGAGCATCTCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3763
425 CCTGAGCATCTCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 484
3764 AGCTGCTAGGCTCTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3823
485 AGCTGCTAGGCTCTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 544
3824 ACTGCACTTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3883
545 ACTGCACTTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 604
3884 ACTGCACTTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3943
605 ACTGCACTTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 664
3944 -AGGACTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4002
665 MAGGACTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 724
4003 AGAGGGGAGCACTGACCTTCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4062
725 AGAGGGGAGCACTGACCTTCTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 784
4063 CATGCCGAGATGCGCATGCGCATGCGCATGCGCATGCGCATGCGCATGCGCATGCGCAT 4122
785 CATGCCGAGATGCGCATGCGCATGCGCATGCGCATGCGCATGCGCATGCGCATGCGCAT 844
4123 ACCGCCAAGCACTTCAACCAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4182
845 ACTTNCACACTTAC-----CAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 899
4183 T 4183

Db 900 T 900

RESULT 8
BUI83160
LOCUS
DEFINITION
BUI83160
AGENCOURT 7969906 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6147102
5', mRNA sequence.
ACCESSION
BUI83160
VERSION
BUI83160.1 GI:22697144
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
NIH-MGC http://mgi.nci.nih.gov/
1 (bases 1 to 861)
NATIONAL INSTITUTES OF HEALTH, MAMMALIAN GENE COLLECTION (MGC)
JOURNAL
Unpublished
AUTHORS
Contact: Robert Strausberg, Ph.D.
Email: c9abs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DRP
COMMENT
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LIML at:
http://image.llnl.gov
Plate: LLM13476 row: g column: 07
High quality sequence stop: 641.
Location/Qualifiers
1. 861
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6147102"
/issue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2 kb. Library constructed by Life
Technologies."

BASE COUNT 215 a 236 c 234 g 175 t 1 others

ORIGIN

Query Match 16.3%; Score 712.6; DB 13; Length 861;
Best Local Similarity 98.2%; Pred. No. 3.7e-174;
Matches 763; Conservative 0; Mismatches 9; Indels 5; Gaps 4;

936 GATGAGCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 995
77 GATGAGCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 136
996 GTGTCCCTCTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1055
137 GTGTCCCTCTCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 196
1056 ACTGCAAGAAACCCCAAGACAGCTTGAAGTGAAGAGGAGGAGGAGGAGGAGGAGGAG 1115
197 ACTGCAAGAAACCCCAAGACAGCTTGAAGTGAAGAGGAGGAGGAGGAGGAGGAGGAG 256
1116 CAAGAAGCCCTCGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1175
257 CAAGAAGCCCTCGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 316
1176 CGTAATTGAGAGCCCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1235
317 CGTAATTGAGAGCCCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 376
1236 GAAAGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1295
377 GAAAGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 436

QY 1296 GAACATGGGACCGTTACCAACCTGCGGAATCTCCGGGAATTCACCTTCCCTGAA 1355
 DB 437 GAACATGGGACCGTTACCAACCTGCGGAATCTCCGGGAATTCACCTTCCCTGAA 496
 QY 1356 GGGCTTCTCTGTTGGCACTGGGAAAGAACCAAGTCTTAATGCTTTTAAACAGGCTGT 1415
 DB 497 GGGCTTCTCTGTTGGCACTGGGAAAGAACCAAGTCTTAATGCTTTTAAACAGGCTGT 556
 QY 1416 GGCAGAGGCGGAGCTGAGATCTACCTGTCAGGCGGAGGCTGTGAGACCTGATGAA 1475
 DB 557 GGCAGAGGCGGAGCTGAGATCTACCTGTCAGGCGGAGGCTGTGAGACCTGATGAA 616
 QY 1476 GAATGACAGAGATACACTTTCAGGAAAGCTGGAATATATACCTGTTTATAGGCGG 1535
 DB 617 GAATGACAGAGATACACTTTCAGGAAAGCTGGAATATATACCTGTTTATAGGCGG 676
 QY 1536 CAATGACCTCTGATGATTTGCAATGATCTGCTCACTTATTCCTCAGAACTTCACAGA 1595
 DB 677 CAATGACCTCTGATGATTTGCAATGATCTGCTCACTTATTCCTCAGAACTTCACAGA 736
 QY 1596 CAACATGGAAGGCGCTGAGCAATCCCTCAATGCT-GAGGTTCTCCGGGCAATTG-TGAC 1653
 DB 737 CAACATGGAAGGCGCTGAGCAATCCCTCAATGCTGAGGTTCTCCGGGCAATTGTTGACC 796
 QY 1654 CTGGTGAAGGCTG--CTTGAGATGCT-CAACCTGAGGAGCTGTACAGAGAAAAA 1707
 DB 797 CTGGTGAAGGCGGCTTGAATGCTCAACCTGAGGAGCTGTACAGAGAAAAA 853

RESULT 9
 BI912651 880 bp mRNA linear EST 16-OCT-2001
 LOCUS 603290988F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:522033 5'
 DEFINITION mRNA sequence.
 BI912651
 BI912651.1 GI:16176846
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgsb@r-mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
 http://image.llnl.gov
 plate: LLAM1553 row: 0 column: 22
 High quality sequence start: 27
 High quality sequence stop: 726.
 Location/Qualifiers
 1. 880
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:522033"
 /issue_type="leukocyte"
 /lab_host="DH10B"
 /clone_lib="NIH MGC_118"
 /note="Vector: PCMV-SPORT6, Site 1: NotI; Site 2: EcoRV
 (destroyed); RNA source leukocytes from anonymous pool of
 non-activated adult donors. Library is oligo-dT primed
 and directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 1.2-3.3 kb. Library is normalized and enriched for

FEATURES

source

1. 880
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:522033"
 /issue_type="leukocyte"
 /lab_host="DH10B"
 /clone_lib="NIH MGC_118"
 /note="Vector: PCMV-SPORT6, Site 1: NotI; Site 2: EcoRV
 (destroyed); RNA source leukocytes from anonymous pool of
 non-activated adult donors. Library is oligo-dT primed
 and directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 1.2-3.3 kb. Library is normalized and enriched for

full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 027. Note:
 this is a NIH MGC Library."

Query Match 16.2%; Score 708.2; DB 12; Length 880;
 Best Local Similarity 98.3%; Pred. No. 5.3e-173;
 Matches 768; Conservative 0; Mismatches 8; Indels 5; Gaps 5;

BASE COUNT 201 a 271 c 228 g 179 t 1 others
 ORIGIN
 QY 2619 TGCAGCCAACTTGTGACCATCTCCGCAATGCTTGGAGCTCTGCATAGAGTGCC 2678
 DB 27 TGCAGCCAACTTGTGACCATCTCCGCAATGCTTGGAGCTCTGCATAGAGTGCC 86
 QY 2679 CAGAGCTGCTGTAACCTGCTGAGCTTCCGAACCCCACTATATGCGGAGGTCTT 2738
 DB 87 CAGAGCTGCTGTAACCTGCTGAGCTTCCGAACCCCACTATATGCGGAGGTCTT 146
 QY 2739 GGGAAACCCAGACAGTGCCTGAGAGGAGGAGGCTTGTGTAATGCTTGTGAC 2798
 DB 147 GGGAAACCCAGACAGTGCCTGAGAGGAGGAGGCTTGTGTAATGCTTGTGAC 206
 QY 2799 CTGCG- GAGAACTCCCAAGACTAGACAGGCTGAGAGGCTTCAAGCCGATACCGGA 2857
 DB 207 CTGCGCGGAGAACTCCCAAGACTAGACAGGCTGAGAGGCTTCAAGCCGATACCGGA 266
 QY 2858 GCAGCATGCGCGAGCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 2917
 DB 267 GCAGCATGCGCGAGCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 326
 QY 2918 TGCTGAGCCCTTCTCCAGAACATCAGCTCCCTGCTGAGGAGTGGCTCCAGATA 2977
 DB 327 TGCTGAGCCCTTCTCCAGAACATCAGCTCCCTGCTGAGGAGTGGCTCCAGATA 386
 QY 2978 GGTCTTCTTCTTGGCCGAGACTGATCACCAGAAATGCAATCCAGCTGGCCA 3037
 DB 387 GGTCTTCTTCTTGGCCGAGACTGATCACCAGAAATGCAATCCAGCTGGCCA 446
 QY 3038 GAGCCCTTGGAGCAATATGCTTGAACCACTTGGAGCAAAAGAGACCTTGAGCTGA 3097
 DB 447 GAGCCCTTGGAGCAATATGCTTGAACCACTTGGAGCAAAAGAGACCTTGAGCTGA 506
 QY 3098 GAGCAGAGATGCCATCACTGCTCCAGTCAAGATGAGCCCTTCTGAGAAACCTTCGGA 3157
 DB 507 GAGCAGAGATGCCATCACTGCTCCAGTCAAGATGAGCCCTTCTGAGAAACCTTCGGA 566
 QY 3158 ATAGTAACCAACGTAATCCCATCAAGCCAGCAATGAGAACTGGGCAAGTCTCTGT 3217
 DB 567 ATAGTAACCAACGTAATCCCATCAAGCCAGCAATGAGAACTGGGCAAGTCTCTGT 626
 QY 3218 GTACAGAGTGAAGGCTTCAATAGTTCGAAC- TGTGTCCAGAGCTTCGAGCAGCA 3276
 DB 627 GTACAGAGTGAAGGCTTCAATAGTTCGAAC- TGTGTCCAGAGCTTCGAGCAGCA 686
 QY 3277 GACATCAAGTGTGGCGGCGGCTGGTGACTCTGACTACA- GCAGTGGAGACTGAGC 3335
 DB 687 GACATCAAGTGTGGCGGCGGCTGGTGACTCTGACTACA- GCAGTGGAGACTGAGC 746
 QY 3336 AAACAATCCAGTACCTACCAATCTTGG- AGGGAGACTT- CTTGAGCATTTGAGGG 3393
 DB 747 CAGCAATCCAGTACCTACCAATCTTGGAGGAGGAGGAGGAGGAGGAGGAGG 806
 QY 3394 G 3394
 DB 807 G 807

RESULT 10
 BX354650 1186 bp mRNA linear EST 05-MAY-2003
 LOCUS BX354650 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 DEFINITION cDNA clone CS0DC033YA08 5-PRIME, mRNA sequence.
 ACCESSION BX354650

```

VERSION      BX34650.1  GI:30379890
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 1186)
              Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
              Full-length cDNA libraries and normalization
              Unpublished
JOURNAL      Contact: Genoscope
              Genoscope - Centre National de Sequencage
              BP 191 91006 Evry cedex - France
              Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
              Library was constructed by Life Technologies, a division of
              Invitrogen. This sequence belongs to sequence cluster 1474.f,
              Contact: Feng Liang Email : fliang@lifetech.com URL :
              http://fulllength.invitrogen.com/Invitrogen Corporation 1600
              Faraday Avenue Genoscope sequence ID : CS0DC023BA04Qp1.
FEATURES
  source
    1..1186
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="CS0DC023YA08"
      /issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
      /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
      /note="1st strand cDNA was primed with a NotI-oligo(dT)
      primer. Five prime end enriched, double-strand cDNA was
      digested with Not I and cloned into the Not I and EcoR V
      sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT   288 a      315 c      281 g      202 t      100 others
ORIGIN
Query Match 15.7%; Score 687.8; DB 13; Length 1186;
Best Local Similarity 91.0%; Pred. No. 1.4e-167;
Matches 711; Conservative 25; Mismatches 42; Indels 3; Gaps 1;

QY 2929 TTCTTCAGAAATCCAGCTCCCTGCTGCGGATGGCTCCAGATACCTCTTCTT 2988
    |||||
DB 65 TTCTTCAGAAATCCAGCTCCCTGCTGCGGATGGCTCCAGATACCTCTTCTT 124

QY 2989 GCCCCAGACTGCATCCACCCAAATTCAGAAATTCACCTCCAGTGGCCAGAGCCCTTTGG 3048
    |||||
DB 125 GCCCCAGACTGCATCCACCCAAATTCAGAAATTCACCTCCAGTGGCCAGAGCCCTTTGG 184

QY 3049 ACCAAATATGCTTGAACCACTTGAAGCAAAACAGAGACCTGACCTGAGAGCAGAGATG 3108
    |||||
DB 185 ACCAAATATGCTTGAACCACTTGAAGCAAAACAGAGACCTGACCTGAGAGCAGAGATG 244

QY 3109 CCCATCACTGTCCCACTCAGAAATGAGCCCTTCTTGAAACCCCTCGGAATATGTAATAC 3168
    |||||
DB 245 CCCATCACTGTCCCACTCAGAAATGAGCCCTTCTTGAAACCCCTCGGAATATGTAATAC 304

QY 3169 ACGTACCCCACTCAAGCAGCAGCATTTGAGAACTGGGGCAGTACTTCTGTGACAGAGTGG 3228
    |||||
DB 305 ACGTACCCCACTCAAGCAGCAGCATTTGAGAACTGGGGCAGTACTTCTGTGACAGAGTGG 364

QY 3229 AAGGCTTCATATGTTTCCAACTCTGTCCACAGAGTCCGACAGCAGACATCAAAATG 3288
    |||||
DB 365 AAGGCTTCATATGTTTCCAACTCTGTCCACAGAGTCCGACAGCAGACATCAAAATG 424

QY 3289 GTGGCCCGCCCTGGGTGACTCTCTGACTTACAGCAGATGGAGTGCACCAAACTCCAGT 3348
    |||||
DB 425 GTGGCCCGCCCTGGGTGACTCTCTGACTTACAGCAGATGGAGTGCACCAAACTCCAGT 484

QY 3349 GACCTACCAACATCTTTGAGGGAGCTCTTTGAGCAGTTGGAGGGATGGGATCTTGAG 3408
    |||||
DB 485 GACCAACCAACATCTTTGAGGGAGCTCTTTGAGCAGTTGGAGGGATGGGATCTTGAG 544

QY 3409 ACTGACACCAACATCTTTGAGGGAGCTCTTTGAGCAGTTGGAGGGATCTTGAG 3465
    |||||
DB 545 ACTGACACCAACATCTTTGAGGGAGCTCTTTGAGCAGTTGGAGGGATCTTGAG 604

```

```

QY 3466 TCTACGACACTGGAGGGAGGACAGCAGACTAATGTGGCAGCGGAGGGCCAGAGCT 3525
    |||||
DB 605 TTTACGACACTGGAGGGAGGACAGCAGACTAATGTGGCAGCGGAGGGCCAGAGCT 664

QY 3526 AGGAGACATGCGACGCCAGCGCTTGGAGCTGTGATAGCGAAATGAAAAACAGCCCGACATC 3585
    |||||
DB 665 AGGAGACATGCGACGCCAGCGCTTGGAGCTGTGATAGCGAAATGAAAAACAGCCCGACATC 724

QY 3586 AACCTGGAGAAAGACTGGAAGCTGTGCACCTTTATTTGGGTCAACGACTGTGCAT 3645
    |||||
DB 725 AATCTGGAGAAAGACTGGAAGCTGTGCACCTTTATTTGGGTCAACGACTGTGCAT 784

QY 3646 TACTGTGAGAAATCCGAGGCGCCACTTGGCCAGCGAAATATGTTGACACATCAACAGGCC 3705
    |||||
DB 785 TACTGTGAGAAATCCGAGGCGCCACTTGGCCAGCGAAATATGTTGACACATCAACAGGCC 844

QY 3706 C 3706
    |||||
DB 845 C 845

RESULT 11
BM549919 1033 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT 6544181 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5745887
DEFINITION 5', mRNA sequence.
ACCESSION BM549919
VERSION BM549919.1 GI:18785712
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1033)
              NIH-MGC http://imgc.nci.nih.gov/.
              National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished
JOURNAL Contact: Robert Strausberg, Ph.D.
              Email: cgabs-remail.nih.gov
              Tissue Procurement: Life Technologies, Inc.
              cDNA Library Preparation: Life Technologies, Inc.
              DNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LIML at:
              http://image.llnl.gov
              Plate: LIML2769 row: 1 column: 24
              High quality sequence stop: 626.
FEATURES
  source
    1..1033
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:5745887"
      /issue_type="leukocyte"
      /lab_host="DH10B"
      /clone_lib="NIH_MGC_118"
      /note="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV
      (destroyed); RNA source leukocytes from anonymous pool of
      non-activated adult donors. Library is oligo-dT primed
      and directionally cloned (EcoRV site is destroyed upon
      cloning). Average insert size 1.7 kb, insert size range
      1.2-3.3 kb. Library is normalized and enriched for
      full-length clones and was constructed by C. Gruber
      (Invitrogen). Research Genetics tracking code 027. Note:
      this is a NIH MGC Library."
BASE COUNT 261 a      292 c      291 g      188 t      1 others
ORIGIN
Query Match 15.5%; Score 677.4; DB 12; Length 1033;
Best Local Similarity 98.5%; Pred. No. 6.4e-165;
Matches 736; Conservative 0; Mismatches 6; Indels 5; Gaps 5;

```



```

QY 3057 GCTTGAACCACTTGAAGCAAAAAGAGACCTGGACCTGAGAGAGATGCCATCAC 3116
DB 153 GCTTGAACCACTTGAAGCAAAAAGAGACCTGGACCTGAGAGAGATGCCATCAC 212
QY 3117 CTGTCCCACTCAGAAATGAGCCCTTCTGAGAACCCCTCGAATATGTAATCACTAC 3176
DB 213 CTGTCCCACTCAGAAATGAGCCCTTCTGAGAACCCCTCGAATATGTAATCACTAC 272
QY 3177 CATCAAGCCGCAATGAGAACTGGGGCACTGACCTCTGTGTAAGAGAGAGAGGCTTC 3236
DB 273 CATCAAGCCGCAATGAGAACTGGGGCACTGACCTCTGTGTAAGAGAGAGGCTTC 332
QY 3237 CAATAGTGTCCAACTCTGTCCACAGCTCCGACAGAGACATCAAGAGTGGCCGC 3296
DB 333 CAATAGTGTCCAACTCTGTCCACAGCTCCGACAGAGACATCAAGAGTGGCCGC 392
QY 3297 CTTGGGTGACTCTCTGACTACAGCAGTGGAGCTCGAACAACTCCAGTGAACCTACC 3356
DB 393 CTTGGGTGACTCTCTGACTACAGCAGTGGAGCTCGAACAACTCCAGTGAACCTACC 452
QY 3357 CACATCTTGAAGGGGACTCTCTTGGAGCATTTGAGGGGATGGGAACTTGGAGACTCAC 3416
DB 453 CACATCTTGAAGGGGACTCTCTTGGAGCATTTGAGGGGATGGGAACTTGGAGACTCAC 512
QY 3417 CACACTGCCCAATCTTGAAGAACTTCAACCTTACCTCTTGGCTTCTCTACAGCAC 3476
DB 513 CACACTGCCCAATCTTGAAGAACTTCAACCTTACCTCTTGGCTTCTCTACAGCAC 572
QY 3477 CTGGAGAGGGGAGACAGAGACTAAATGTGGAGCGGAAAGGGGCAAGACTTACATGCC 3536
DB 573 CTGGAGAGGGGAGACAGAGACTAAATGTGGAGCGGAAAGGGGCAAGACTTACATGCC 632
QY 3537 AGCCCAAGGCTTGGAGACTCTGTGTAAGCGAATGAAAGAAACAGCCCGACATCACTGAG 3594
DB 633 AGCCCAAGGCTTGGAGACTCTGTGTAAGCGAATGAAAGAAACAGCCCGACATCACTGAG 692
QY 3595 AAAGACTGGAAGCTGGTCACTCTTATGGGGTCAAGAGCTTGTGTCTTACTGAG 3654
DB 693 AAAGACTGGAAGCTGGTCACTCTTATGGGGTCAAGAGCTTGTGTCTTACTGAG 752
QY 3655 AATCCGAGAGCCCACTTGGCCAGGATATGTTTCAAGACATCCAAAGCCCTGAGCATC 3714
DB 753 AATCCGAGAGCCCACTTGGCCAGGATATGTTTCAAGACATCCAAAGCCCTGAGCATC 812
QY 3715 CTCTCTGAGAGAGCTCCCAAGGGCTTCTGTAAGCTGAGAGCTGAGAGCTGAGCT 3771
DB 813 CTCTCTGAGAGAGCTCCCAAGGGCTTCTGTAAGCTGAGAGCTGAGAGCTGAGCT 872
QY 3772 AGCCTGTACAGAGGCGCAAGGCGGGGAA 3798
DB 873 AGCCTGTACAGAGGCGCAAGGCGGGGAA 899

```

```

RESULT 12
LOCUS B1912073 763 bp mRNA linear EST 16-OCT-2001
DEFINITION 603068631F1 NIH_MGC_118 Homo sapiens cDNA clone IMAGE:5217606 5',
mRNA sequence.
VERSION B1912073
KEYWORDS B1912073.1 GI:16176045
SOURCE EST.
ORGANISM Homo sapiens (human)
Homo sapiens
Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 763)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.

```

```

FEATURES
source
1..763
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5217606"
/issue_type="leukocyte"
/lab_host="DH10B"
/clone_lib="NIH_MGC_118"
/note="Vector: pCMV-Sport6; Site 1: NotI; Site 2: EcoRV
(destroyed); RNA source leukocytes from anonymous pool of
non-activated adult donors. Library is oligo-dT primed
and directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.7 kb, insert size range
1.2-3.3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 027. Note:
this is a NIH MGC Library."
BASE COUNT 177 a 238 c 193 g 155 t
ORIGIN
Query Match 15.3%; Score 671.4; DB 12; Length 763;
Best Local Similarity 95.9%; Pred. No. 1.9e-163;
Matches 732; Conservative 0; Mismatches 26; Indels 5; Gaps 4;
QY 2612 TGTATTCTGACGCAACTTTGTTGACATCTCCGCAATGCTTGGAGCTCTGCATAGAG 2671
DB 2 TGTATTCTGACGCAACTTTGTTGACATCTCCGCAATGCTTGGAGCTCTGCATAGAG 60
QY 2672 AGGTGCGCAGAGCTCGTGCAACCTGTGAGCTTCTGAAACCCCAATCATGCGGAGG 2731
DB 61 AGGTGCGCAGAGCTCGTGCAACCTGTGAGCTTCTGAAACCCCAATCATGCGGAGG 120
QY 2732 TGTTCCTGGGAAACCCAGACAGTGGCCAGTGAGAGGCGAGCTTTGTGTAATGCG 2791
DB 121 TGTTCCTGGGAAACCCAGACAGTGGCCAGTGAGAGGCGAGCTTTGTGTAATGCG 179
QY 2792 TTCTGACCTTGGGAGAACTCCCAAGAGCTGAGGCTTGAAGGCTTGAAGGCTT 2851
DB 180 TTCTGACCTTGGGAGAACTCCCAAGAGCTGAGGCTTGAAGGCTTGAAGGCTT 239
QY 2852 ACCGAGAGAGCATGCGGAGAGCTGTGGGTGAGGCGGCTATGACAGAGAGAGATTTCT 2911
DB 240 ACCGAGAGAGCATGCGGAGAGCTGTGGGTGAGGCGGCTATGACAGAGAGAGATTTCT 299
QY 2912 CTGTGTGCTGACAGCCCTTCTTCAGAACTCCAGCTCCCTGTCTGGCGAGTGGCTCC 2971
DB 300 CTGTGTGCTGACAGCCCTTCTTCAGAACTCCAGCTCCCTGTCTGGCGAGTGGCTCC 359
QY 2972 CAGATACGTCCTTCTTTGGCCAGAGCTGACCAATTCAGAAATTCCTCCAGC 3031
DB 360 CAGATACGTCCTTCTTTGGCCAGAGCTGACCAATTCAGAAATTCCTCCAGC 419
QY 3032 TGGCCAGAGCCCTTGGAGCAAA--TATGCTTGAACCACTTGGAAAGCAAGAGACCT 3089
DB 420 TGGCCAGAGCCCTTGGAGCAAAATATGTAACCACTTGAAGCAAGAGAGACCTT 479
QY 3090 GGAAGCTGAGAGAGAGATGCCATCACTGTCCCATCTCAGATGAGACCTTCTGAGAAC 3149
DB 480 GGAAGCTGAGAGAGAGATGCCATCACTGTCCCATCTCAGATGAGACCTTCTGAGAAC 539
QY 3150 CCCTGGAAATGTAATCACTGACCTTCAAGAGGCGCATTTGAGAGTGGGCGAGTGA 3209
DB 540 CCCTGGAAATGTAATCACTGACCTTCAAGAGGCGCATTTGAGAGTGGGCGAGTGA 599

```


QY 3210 CTTCCTGTGTACAGAGTGAAGGCTTCCAAATAGTGTCCAAAC-CTGTGTCCACAGAGCTCC 3268
 DB 600 CTTCCTGTGTACAGAGTGAAGGCTTCCAAATAGTGTCCAAACATCTGTGTCCACAGAGCTCC 659
 QY 3269 GACCAGCAGACATCAAAAGTGTGGCCGCTGGGTGACTCTGTGTACATCAAGAGTGGAG 3328
 DB 660 GACCAGCAGACCTCAAAAGTGTGGCCGCTGGGTGACTCTGTGTACATCAAGAGTGGAG 719
 QY 3329 CTGACCAACAACACTCCAGTACCTACCCACATCTTGGAGGGG 3371
 DB 720 CTGACCAACAACCTCCAGTACCTACCCACATCTTGGAGGGG 762

RESULT 13
 BI488437 823 bp mRNA linear EST 28-AUG-2001
 LOCUS 603020936F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5191712 5',
 DEFINITION mRNA sequence.
 ACCESSION BI488437
 VERSION BI488437.1 GI:15327665
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 823)
 NIH-MGC http://mgs.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Sequencing by: Incyte Genomics, Inc.
 DNA Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 plate: LLAM1479 row: g column: 09
 High quality sequence, stop: 808.
 Location/Qualifiers
 1. 823
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5191712"
 /lab_host="DH10B"
 /clone_id="NIH_MGC_114"
 /note="Organ: brain; Vector: pCMV-Sport6; Site 1: NotI;
 Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
 male brains, age range 23-27 yo. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.5 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 019. Note:
 this is a NIH_MGC Library."
 BASE COUNT 213 a 239 c 214 g 157 t
 ORIGIN
 Query Match 14.3%; Score 627.8; DB 12; Length 823;
 Best Local Similarity 97.2%; Pred. No. 4.7e-152;
 Matches 660; Conservative 0; Mismatches 17; Indels 2; Gaps 2;

DB 266 CATCAAGCCAGCCATTGAGAACTGGGGCAGTGACTTCTGTGTACAGAGTGAAGGCTTC 325
 QY 3237 CAATAGTGTTCACACTCTGTGTCCACAGCTCCGACCAAGACATCAATAGTGTGGCCG 3296
 DB 326 CAATAGTGTTCACACTCTGTGTCCACAGCTCCGACCAAGACATCAATAGTGTGGCCG 385
 QY 3297 CTTGGTGAATCTCTGACTACAGCAGTGGAGGTGACCAAAACATCTCCAGTACCTACC 3356
 DB 386 CTTGGTGAATCTCTGACTACAGCAGTGGAGGTGACCAAAACATCTCCAGTACCTACC 445
 QY 3357 CACATCTTGGAGGGGACTCTCTTGGAGCATTTGAGGGGATGGAACTTGGAGACTCAC 3416
 DB 446 CACATCTTGGAGGGGACTCTCTTGGAGCATTTGAGGGGATGGAACTTGGAGACTCAC 505
 QY 3417 CACACTGCCCAACATTTCTGAAGAAGTTCAACCTTACCTCTTGGCTTCTCTACAGAC 3476
 DB 506 CACACTGCCCAACATTTCTGAAGAAGTTCAACCTTACCTCTTGGCTTCTCTACAGAC 565
 QY 3477 CTGGAGAGGGGAGCAGAGACTTAATGTGGAGGGGAGGGGAGGAGCTAGGAGCATGCC 3536
 DB 566 CTGGAGAGGGGAGCAGAGACTTAATGTGGAGGGGAGGGGAGGAGCTAGGAGCATGCC 625
 QY 3537 AGCCCAAGGCTTGGAGCTGTGTAGAGCGAATGAAGAAACAGCCCGACATCAACCTGAGAA 3596
 DB 626 AGCCCAAGGCTTGGAGCTGTGTAGAGCGAATGAAGAAACAGCCCGACATCAACCTGAGAA 685
 QY 3597 AGACTGAGAGCTGTGTACACTTTC-ATTGGGGTCAACGACTTGTCTATTACTGTGAGA 3655
 DB 686 AGACTGAGAGCTGTGTACACTTTCATTGAGGGTCAACGACTTGTCTATTACTGTGAGA 745
 QY 3656 ATCCGAGGCGCCACTTGGCCAGGAAATGTGTGAGACATCCAAAGGCGCTGAGATCC 3715
 DB 746 ATCCGAGGCGCCACTTGGCCAGGAAATGTGTGAGACATCCAAAGGCGCTGAGATCC 804
 QY 3716 TCTCTGAGAGCTGCCAAG 3724
 DB 805 TCTCTGAGAGCTGCCAAG 823

RESULT 14
 BM714799 607 bp mRNA linear EST 28-FEB-2002
 LOCUS BM714799
 DEFINITION UI-E-EJ0-ahs-f-21-0-UI-r1 UI-E-EJ0 Homo sapiens cDNA clone
 UI-E-EJ0-ahs-f-21-0-UI 5', mRNA sequence.
 ACCESSION BM714799
 VERSION BM714799.1 GI:19028057
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 607)
 Bonaldo, M.F., Lennon, G. and Soares, M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 JOURNAL
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 8250
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 Reverse.

FEATURES		Location/Qualifiers
Source	1..607	/organism="Homo sapiens"
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="UI-E-EJ0-ahs-f-21-0-UI"	
	/tissue_type="fetal eyes, lens, eye anterior segment, optic nerve, retina, Retina Foveal and Macular, RPE and Choroid"	
	/dev_stage="fetal and adult"	
	/lab_host="DH10B (Life Technologies) (T1 phage resistant)"	
	/clone_1b="UI-E-EJ0"	
	/note="Organ: eye; Vector: pRT73-Pac (Pharmacia) with a modified polylinker; Site 1: Ecor I; Site 2: Not I; UI-E-EJ0 is a subcloned cDNA library constructed according to Bonaldo, Lemmon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an Ecor I adaptor, digested with Not I, and cloned directionally into pRT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dr)18 tail. The sequence tags for this library are: fetal eyes, AGAATCAGA; lens, CGATTAGCGA; eye anterior segment, AATGCCGAT; optic nerve, CCATTAGTGC; retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, ACCCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."	
	BASE COUNT	156 a 178 c 156 g 117 t
ORIGIN		
Query Match		13.9%; Score 607; DB 12; Length 607;
Best Local Similarity		100.0%; Pred. No. 9.9e-147;
Matches 607; Conservative		0; Mismatches 0; Indels 0; Gaps 0;
Qy	3529	GACATGCCAGCCAGCGCTGGACCTGTGAGAGGAATGAAAAACAGCCCGACATCAAC 3588
Db	1	GACATGCCAGCCAGCGCTGGACCTGTGAGAGGAATGAAAAACAGCCCGACATCAAC 60
Qy	3589	CTGGAGAAAGACTGGAGAGCTGTCACTTTCAATGGGGTCAACGACTTGTCTATTAC 3648
Db	61	CTGGAGAAAGACTGGAGAGCTGTCACTTTCAATGGGGTCAACGACTTGTCTATTAC 120
Qy	3649	TGTGAGAAATCCGAGGCGCCACTTGGCCAGGAATATTTTTCAGCATCCACAGGCCCTTG 3708
Db	121	TGTGAGAAATCCGAGGCGCCACTTGGCCAGGAATATTTTTCAGCATCCACAGGCCCTTG 180
Qy	3709	GACATCCTCTCTGAGAGAGCTCCCAAGGGCTTTTGTCAACGTGGGAGGATCATGAGCTG 3768
Db	181	GACATCCTCTCTGAGAGAGCTCCCAAGGGCTTTTGTCAACGTGGGAGGATCATGAGCTG 240
Qy	3769	GCTAGCTGTACCAAGGGCCAAAGCGGGAATGTGCCATGCTGGCAGCTCAGAACATGCG 3828
Db	241	GCTAGCTGTACCAAGGGCCAAAGCGGGAATGTGCCATGCTGGCAGCTCAGAACATGCG 300
Qy	3829	ACTTGGCTCAGACACTGCGAAAGCTCCTCTGAGAGAGCAAACTGAAAGAAAGTGAAGTGG 3888
Db	301	ACTTGGCTCAGACACTGCGAAAGCTCCTCTGAGAGAGCAAACTGAAAGAAAGTGAAGTGG 360
Qy	3889	AACCTCCAGAGATGCAATCTCAGTTTCTCTACTGCGACCAATACACAGCGGTGAGGAC 3948
Db	361	AACCTCCAGAGATGCAATCTCAGTTTCTCTACTGCGACCAATACACAGCGGTGAGGAC 420
Qy	3949	TTTGGGTTTGTGTGAGAGCTTTTCTTCCAAACAACACTCAACCCCACTGAGAGAGAGG 4008
Db	421	TTTGGGTTTGTGTGAGAGCTTTTCTTCCAAACAACACTCAACCCCACTGAGAGAGAGG 480
Qy	4009	GACACTGACCTCACTTCTTCTCCGAGAGACTGTTTTCATCTTCAGACCGCGGAGATGCC 4068
Db	481	GACACTGACCTCACTTCTTCTCCGAGAGACTGTTTTCATCTTCAGACCGCGGAGATGCC 540
Qy	4069	GAGATGGCATTGCGACTGTGGAACAACATGCTGGAACCAATGGGCGCGCAAGACTACTCC 4128

Db	541	GAGATGGCATTGCGACTGTGGAACAACATGCTGGAACCAATGGGCGCGCAAGACTACTCC 600
Qy	4129	AACACT 4135
Db	601	AACACT 607
RESULT 15		
AK015993		
LOCUS		
DEFINITION		AK015993 1562 bp mRNA linear HTC 05-DEC-2002
ACCESSION		Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4930539A06 product:weakly similar to PHOSPHOLIPASE B [Rattus norvegicus], full insert sequence.
VERSION		AK015993.1 GI:12854560
KEYWORDS		HTC; CAP trapper.
SOURCE		Mus musculus (house mouse)
ORGANISM		Mus musculus
REFERENCE		Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS		1 Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)
TITLE		2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL		3 Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Iizawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE		20499374
PUBMED		11042159
REFERENCE		4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Iizawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Giesi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsumoto, Y., Nishino, T., Pesole, G., Quackenbush, J., Schriml, L. M., Staudt, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Balderelli, R., Barsch, G., Blake, J., Boffelli, D., Bojunga, N., Carinci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Guelincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Maehima, J., Mazzarelli, J., Mombaur, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schombach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohseuki, S. and Hayashizaki, Y. Functional annotation of a full-length mouse cDNA collection Nature 409 (6821), 685-690 (2001)
JOURNAL		21085660
MEDLINE		11217851
PUBMED		5
REFERENCE		


```

QY 1033 AGAAACGCACTACTGACCAGACTGCAGAAACCCCAAGCAAGCTTGAAGTAAGAGAA 1092
Db 1236 AAAAATAGCACTACTGAGCCAGACGACTGAACCTATTAACAAGCTTCAGATGAAGAA 1295
QY 1093 GGACCGGAATCAGATGTCTCTGACAAAGACCCCTCCGATAGGGTTCCOACCTCAGTTCA 1152
Db 1296 GGATCAGAAATTCACCTGTCTGACAGAACCCCTCGAACTCAATCCCAACACAGGTACA 1355
QY 1153 AGGCTG 1158
Db 1356 TGGCTG 1361

```

Search completed: January 6, 2004, 14:41:56
 Job time : 8450 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 6, 2004, 18:56:30 ; Search time 27 Seconds

(Without alignments)
2539.442 Million cell updates/sec

Title: US-10-054-691-2

Sequence: 1 MGIRPGIFLELLILGQGT.....RCRRGRREDPPMSLRVAL 1458

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database: SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query #	Length	DB ID	Description
1	5987.5	77.1	1458	1	PHLX_RABIT
2	149	1.9	1458	1	YAMB_SCHPO
3	138	1.8	1455	1	FACA_HUMAN
4	131.5	1.7	1361	1	GLI4_XENTLA
5	128	1.6	1155	1	C2TA_MOUSE
6	126.5	1.6	1534	1	YFAS_ECOLI
7	126.5	1.6	6885	1	SNE2_HUMAN
8	125.5	1.6	928	1	ATC1_YARLI
9	125	1.6	1178	1	YFAS_ECO57
10	123.5	1.6	1534	1	YFAS_ECO57
11	123.5	1.6	2672	1	GCN1_YEAST
12	123	1.6	1039	1	PDP2_ARATH
13	122.5	1.6	713	1	GALK_SCHPO
14	122	1.6	1811	1	SC71_SCHPO
15	121.5	1.6	1097	1	TOLR_DROME
16	121.5	1.6	8797	1	SNE1_HUMAN
17	121	1.6	953	1	ATC3_BOVIN
18	121	1.6	4725	1	DYHC_DICDI
19	120.5	1.6	2843	1	APC_HUMAN
20	119.5	1.5	1192	1	RTN4_HUMAN
21	119.5	1.5	1276	1	BXD_CLOBO
22	117.5	1.5	1279	1	BCSG_PSEFL
23	116.5	1.5	1302	1	UBPY_CAREL
24	116.5	1.5	2167	1	BEM2_YEAST
25	116.5	1.5	2469	1	TEGU_HSVSA
26	116	1.5	919	1	ATC1_HUMAN
27	116	1.5	1420	1	SRB9_YEAST
28	116	1.5	1489	1	CC25_YEAST
29	115.5	1.5	555	1	G6P1_BUCAP
30	115.5	1.5	1849	1	BIG1_HUMAN
31	115.5	1.5	5596	1	MDN1_HUMAN
32	115	1.5	1301	1	SAC3_YEAST
33	114.5	1.5	1498	1	Y1A9_CLOAB

34	114.5	1.5	2363	1	SPOC_MOUSE
35	114	1.5	919	1	ATC1_RAT
36	114	1.5	2738	1	PGCV_RAT
37	114	1.5	4377	1	ANK3_HUMAN
38	113.5	1.5	880	1	IMB2_HUMAN
39	113.5	1.5	1770	1	R1S1_YEAST
40	113	1.5	879	1	YE07_MYCPN
41	113	1.5	954	1	DRP2_HUMAN
42	113	1.5	1489	1	RNG2_SCHPO
43	113	1.5	1533	1	Y566_HUMAN
44	113	1.5	1932	1	PAB1_SCHPO
45	112.5	1.4	882	1	DIS1_SCHPO

ALIGNMENTS

RESULT 1	ID	PHLX_RABIT	STANDARD	PRT: 1458 AA.
AC	005017			
DT	01-FEB-1994 (Rel. 28, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	Phospholipase A2B1-B precursor (EC 3.1.1.-)			
OS	Oryctolagus cuniculus (Rabbit)			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.			
OX	NCBI_TaxID=9986;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.			
RC	TISSUE=Intestine;			
RX	MEDLINE=93286138; PubMed=8509424;			
RA	Boll W., Schmid-Chanda T., Semenza G., Mantel N.;			
RT	Isolation of cognate cDNAs and characterization of a novel brush			
RT	border protein with esterase and phospholipase activity."			
RL	J. Biol. Chem. 268:12901-12911(1993)			
CC	- FUNCTION: HAS ESTERASE AND PHOSPHOLIPASE A/PHOSPHOLIPASE			
CC	ACTIVITY. CAN CONVERT PHOSPHATIDYLCHOLINE TO FATTY ACIDS AND			
CC	GLYCEROPHOSPHOCHOLINE. COULD BE INVOLVED IN UPTAKE OF DIETARY			
CC	LIPIDS, POSSIBLY INCLUDING LONG CHAIN RETINYL ESTERS.			
CC	- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. BRUSH BORDER.			
CC	- TISSUE SPECIFICITY: INTESTINE.			
CC	- DEVELOPMENTAL STAGE: EXPRESSED IN THE INTESTINE OF ADULT BUT NOT			
CC	BABY RABBIT.			
CC	- SIMILARITY: BELONGS TO THE "GDSL" FAMILY OF LIPOLYTIC ENZYMES.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
DR	EMBL, Z12841; CAI78303.1; ..			
DR	PIR, A45665; A45665.			
DR	InterPro, IPR001087; Lipase_GDSL.			
DR	Pfam, PF00657; Lipase_GDSL.3.			
DR	PROSITE, PS01098; LIPASE_GDSL_SER; 2.			
KW	Hydrolase; Repeat; Signal; Transmembrane.			
FT	SIGNAL	1	19	POTENTIAL.
FT	CHAIN	20	1458	PHOSPHOLIPASE A2B1-B.
FT	DOMAIN	20	1415	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	1416	1439	POTENTIAL.
FT	DOMAIN	1440	1458	CYTOSOLIC (POTENTIAL).
FT	DOMAIN	39	1403	4 X 308-326 AA APPROXIMATE REPEATS.
FT	REPEAT	39	347	1.
FT	REPEAT	362	707	2.
FT	REPEAT	708	1054	3.
FT	REPEAT	1064	1403	4.
FT	ACT_SITE	400	400	BY SIMILARITY.

FT ACT SITE 747 747 BY SIMILARITY.
 FT ACT SITE 1103 1103 BY SIMILARITY.
 SQ SEQUENCE 1458 AA, 161343 MW, 901608FP47B1062E6 CRC64;
 Query Match 77.1%; Score 5987.5; DB 1; Length 1458;
 Best Local Similarity 77.6%; Pred. No. 0;
 Matches 1130; Conservative 127; Mismatches 197; Indels 3; Gaps 3;

QY 1 MGLRPGFLELLLLGQCTPQIHTSPKSTLSEGLMPETLKSPFCNPNKLCVMMPSK 60
 DB 1 MAMPSPFLLGLPLRLGRGADQIQTSSGKNTLSEGLMPESLKTFFPCDPTLAESVPS 60
 QY 61 SVHSLSKSDIKFVAALGNLEIPDPGCTGDEKODMTE-REPOVOMGMYTSLDIRFSP 119
 DB 61 SVHSLSRSDIKFVAALGNVETAPDSGADLLEBDCTEKREBQACMGVYTLSDIIGFSP 120
 QY 120 SVMPVCHTKRVI PHDAEDLMTQAQELVNNKENIQLDFQDWKLINVFPSNASQCYL 179
 DB 121 SALMPLC-PETRLVPRGGAEDLMWQATELVRSKRENEQLDFEHDMKLIINVFSTSQCFP 179
 QY 180 CPBAQONGLAGVDELMTGVLDTYQGVPRAPVNLVLSVAVSRQYHGTWLSRAEPC 239
 DB 180 CPBAQOGLVGGMDKLTRLDYIQGVPRAPVNLVLSLAESRRQQAQLSPAEP 239
 QY 240 NCSEETRLAKVVMQMSYQEAANSLASRYSEQSEFTVVFQPFYETPSLSEDPRLQ 239
 DB 240 RCLRETSQTLKVLQNSYLEAMDSILASSKYNQESFAVVFQPFYESSLSALLAEPLQ 239
 QY 300 DSTTLAMHLNNRMEPAGKEDPLSVHGRPMKCPSESPYLSYNSNYLTRLOKPDK 359
 DB 300 DPTTLALSTNNRMEPIGRKEEPFSEKERKPCPTQESPYLFTYRNSQGLTRVSPQOK 359
 QY 360 LEVREGAETRCPPKDPSTVPTSHRLKPADINVIQGLSTRAGNAGSTPGVNLVLT 419
 DB 360 LEVREGELRCPPDPSTVPTSHRLKPADIKIYIGAGMDSLTAGNAGSQGPGLDLVLT 419
 QY 420 QYRGLSNVSGDENIGVTTLANTIREFNSLKGFSVGTGKETSPPNAFLQAVAGRAED 479
 DB 420 QYRGLSNVSGDDNISVTTLANTIREFNSLQGSFSTGTERITTSQAFRQAVAGRAADG 479
 QY 480 LPVOARLVLMKNDTRIHFQEDWKIITLFIQGNLDCDFNDLVHVSPOFTDNIGKALD 539
 DB 480 LIPOAGRLVALMKNDTRIHFQEDWKIITVFIQGNLDCDFNDLPYRSPQFTDNIGKALD 539
 QY 540 ILHAEPRAVNIYVYELYNLRELVOEKYVCPRMILRSLCPVULKPDNSTELALIE 599
 DB 540 ILHAELPRAVNIYVYELISKRELVOETKVSCEPMILRSLCPVULKPDNSTELALIE 599
 QY 600 FNKKFOEKTQLIESGRYDREDFTVVVOPFEFVNDMPKTSSEGLPDNSFPAPDCFHSSK 659
 DB 600 TIKEYGERTQQLIDSGYDTRDFTVVLQPFPEFVNMKPTQDGLPDNSFPAPDCFHSSK 659
 QY 660 SHSRASALMNNMLEPVQKTRHKKFNKINITCPNOVOPFLRTYKNSMOGHGTWLP 719
 DB 660 AHAHAASALMNNMLEPVQKTRHNDFECAVNTCPNQVWPFSLTYKNSVQGFMTWLP 719
 QY 720 RAASALHPTSVHMLRPADIDVVAALGSLTAGNGISGKEDDLPTVTOYGLSYSGAGDG 779
 DB 720 RSPSASAPTSVHMLRPADIDVVAALGSLTAGIGISGKEDDLPTVTOYGLSYSGAGDG 779
 QY 780 SLENVTTLPLILREPNLITGVANGTGDANDTNAFLQAVGAKAEDLMOVOTLMQMK 839
 DB 780 SLDNVTTLPILKOFNSNLGFAVGTDASGTAFFQAVGAKAKADLMQVOTLMQMK 839
 QY 840 DHRHVNFEHDMKYITVLIGSDLCDCYCTDSNLSAANFVHDLRNALDVLHREPVVLNL 899
 DB 840 DHRHVNFEQEDMKYITVIGASDLCDCYCTDSNLSAANFVHDLRNALDVLHREPVVLNL 899
 QY 900 VDLNLPITMQVFLGNPKCPVOQASVLCVCLTRNSQELARLEAFSAIYSSKRELV 959
 DB 900 VDFMNSBVTQVFLGNPKCPVOQASVLCVCLTRNSQELARLEALTAQYOSSLRELV 959
 QY 960 GSGRYDPTQEDFSVVLQPFQNIQLPVVLADGLPDTSPFAPOCIHNNQGFHSQALRALMTNM 1019

DB 960 ESGRYDPTQEDFSVVLQPFHSHIQLPVLDGRLDTSFPAPDCVAPNCKFHSQSLRALMRNM 1019
 QY 1020 LEPGLSKTETLDIRAEMPITTCPTQNEPFLTPKNSNTYPIKAIENWGSDFLCTEKAS 1079
 DB 1020 LEPGLSKTETLDIRAEMPITTCPTQNEPFLTPKNSNTYPIKAIENWGSDFLCTEKAS 1079
 QY 1080 NSVPTSVHOLRPADIRYVAALGSLTTAVGARNNSSDPTSRGSLSMSIGCGDNLETHT 1139
 DB 1080 RGVPSVHOLRPADIRYVAALGSLTTAVGARNNSSDPTSRGSLSMSIGCGDNLETHT 1139
 QY 1140 TLPLNLLKKNPVLGEGSTSTWEGTAGLNYAEGARARMDPAQAMDVERKNSPDINLEK 1199
 DB 1140 TLPLNLLKKNPVLGEGSTSTWEGTAGLNYAEGARARMDPAQAMDVERKNSPDINLEK 1199
 QY 1200 DMKLVTLFVGNDLCHFCENPBGSSBGEVYQHQAQALDVLYEELPRTFVVVVEMLAGL 1259
 DB 1200 DMKLVTLFVGNDLCHFCENPBGSSBGEVYQHQAQALDVLYEELPRTFVVVVEMLAGL 1259
 QY 1260 YOGGKGKCA-MLAONNCTLRHSSSLKQELKKNMNLQHGISFSYWHQYTOREDFA 1318
 DB 1260 HODQGGKCATLLAOSHCTCFKTSQSSVEMQELKKNMNLQSGLSRLSYHQAQEDFA 1319
 QY 1319 VVVQPFQNTLTPLNREGDTLTFPSBDCPHFSDRGHAEMALMNNMLEPVGRKTTSN 1378
 DB 1320 VVVQPFQNTLTPLNREGDTLTFPSBDCPHFSDRGHAEMALMNNMLEPVGRKTTSN 1379
 QY 1379 FTHSRALCKPSPESPLYTLRNSRLPPOAEAPVLYVAVVAAGVGLVGIIGTVW 1438
 DB 1380 FTHSRALCKPSPESPLYTLRNSRLPPOAEAPVLYVAVVAAGVGLVGIIGTVW 1439
 QY 1439 RCRRGGRREDPMSLRT 1455
 DB 1440 RCRRGGRREDPMSLRT 1456

RESULT 2
 YAMB_SCHPO
 ID YAMB_SCHPO STANDARD; PRT; 3655 AA.
 AC Q10064;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein C1F5.11c in chromosome I.
 GN SPAC1F5.11C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Scouras J., Peat N., Hayes J., Baker S., Baaham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltham T., Frazer A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitz E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Rabinowitz E.,
 RA Welford I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gebel C., Fuchs M., Fittz C., Holzer B., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehach H., Reinhardt R., Pohl T.M.,
 RA Beyer P., Zimmermann W., Wedler H., Wambut R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaire V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G., Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J., Dominguez A., Revuelta J., Moreno S., Armstrong J., Forsburg S.L., Carrutti L., Lowe T., McComb W.R., Paulsen I., Potashkin J., Shpatovskii G.V., Usery D., Barrell B.G., Nurse P., "The genome sequence of Schizosaccharomyces pombe.", Nature 415:871-880(2002).

CC -1 SIMILARITY: STRONG, TO YEAST YHR099W.
 CC -1 SIMILARITY: BELONGS TO THE PI3/P14-KINASES FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/announcements or send an email to license@ebi.ac.uk).

DR EMBL: Z68136; CA92239.1; -
 DR PIR: T38084; T38084.
 DR GeneDB: SPombe; SPAC1F5.11c; -
 DR InterPro: IPR003151; PAT.
 DR InterPro: IPR003152; PATC.
 DR InterPro: IPR004403; PI3_P14_kinase.
 DR InterPro: IPR001440; TPR.
 DR Pfam: PF02259; FAT; 1.
 DR Pfam: PF02260; FATC; 1.
 DR Pfam: PF00454; PI3_P14_kinase; 1.
 DR SMART: SM00446; PI3Kc_1.
 DR PROSITE: PS50290; PI3_4_KINASE_3; 1.
 DR Hypothetical protein; Transferrase; Kinase.
 FT DOMAIN 3324 3655 PI3K/PI4K (BY SIMILARITY).
 SQ SEQUENCE 3655 AA; 420774 MW; 50475E3F3FC124A CRC64;

Query Match 1.9%; Score 149; DB 1; Length 3655;
 Best Local Similarity 20.0%; Pred. No. 0.76;
 Matches 167; Conservative 112; Mismatches 307; Indels 248; Gaps 40;

QY 672 MLEPVGQKTR--HKFNKINITCNQVOPFLRTKNSM-----OGHWLQ 717
 DB 367 LURPLAFSLADLHHVRDELNT---QIRKSMITSTNHDLTLGLOTWAKRLILNM 423
 QY 718 RDR--APSALHPTSHALRPADIOVVAALGSLTANGIGSKPDDLPTVTTQYRGLSYS 774
 DB 424 VVRMISLPsi-----PDALFILSLIFDSFV-----NKFSLSNLSLQDFKKKYE 467
 QY 775 AGDGSLENTVTLPLNLRPNRNLTGAVGTGANDTNALNQAAPCAKEDIMSOVQL 834
 DB 468 -----EIKETKSPTRSSPRDLSFSTSVND-----GSFL-----FKNIMFGLRAL 508
 QY 835 MCKMKDHRVNFHEDKVIIVLIGSDLDCTDSNLYSANFVDLRLNADVLHREVR 894
 DB 509 MGLR-----TCKSRCTIEGSGQSGFLTNIKPFELVVF-----QKLPFEVK 551
 QY 895 VLVNVLNPTIMRQVFLGNPDKCPVOQASVLCNVLTLRENSQELARLEAFSRYSRS 954
 DB 552 GF---SYRRP---EQVILETFEC---EESLDRAPISTLPRKDKDCLVAVATIF---599
 QY 955 MEVLVSGRYTQEDFSVVLQPFQNIQLPVLADGLPD---TSFADDCIHPNQ---KTH 1008
 DB 600 -----IHLPSIFLKVETN---LPTFPDQKKNLTLPHIQFLSNSTSSKFL 646
 QY 1009 SOLARALMTNMLEPGSKTETLDLRAMPi-----TCPQNEPFLR-----1049
 DB 647 NTLRLPL-LSRIEELSS---DIRGSVILRLFRLSFTVVSFEATENEPVLRPYSEII 701
 QY 1050 -----TPRNSNTYPIKPAIENMGSDFLCTEKMANSVPTSHQRLPADIKVVAAL 1100
 DB 702 VKCMKLAPNSANSINYYILRALFRQIGGRFESLYK---EWVPLHALLLEANSL---755
 QY 1101 GDSLTVAAGRPNNSSDLPTS-----WRGLSNISGGDGNLEF---1137
 DB 756 -----ISARTPKERKDLFTLCLTIPVRLSLPYMSYLMRPLVMSLXSSQELVQGLR 808

QY 1138 --HTTLPNLIKRF-----NPYLGFESTSTW-----EGTAGLVAAEGARADM 1178
 DB 809 TPELCIDNLTTPPLDIPMAPYIEDLNALMASHQLPYPYNNISHTKLIGKGGNRKL 868
 QY 1179 PAQAMDVERKMSQSPDINEKQWLVLPFGV--DLCHYCENPEALATEVYQIQQAL 1236
 DB 869 ---LDRVQSLKNSPEPN--NDFTLLSISIKGVQPOLH-----YQYVDEAV 910
 QY 1237 DILSELPRAPFNVV---WELASLYQGCGKAMLAQNNCTCLHRSOSLEKQL 1291
 DB 911 NLLSS--PSSDLEVKQQAFTYCNISKLTVYSDATNSLASIRTAADKISKNDFRRP 968
 QY 1292 KKVNNMLQGISFSFYWHQYTORDEFAVVQF---FFONTLTPINERGDTLTFEFS- 1345
 DB 969 YSVIPRMNMGRRSFT---QLSDSDSDTITILASATYGLFATYVDELRE---EAYFWLEK 1021
 QY 1346 -----DCFHPSD---RGHAEMALALNNMLPEVGRKTSNNFTHSRAKLC 1388
 DB 1022 LAVNVIWHDIFYAFVDVIGNHSKFTTNLQKEYI-----ISPHYFAHCLSEVVC 1069

RESULT 3
 ID FACA HUMAN STANDARD; PRT; 1455 AA.
 AC 015360; 075266; 092497; 096118; 09UEA5; 09UEB8; 09UEB9; 09UPK3;
 AC 09Y6M2;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Fancconi anemia group A protein (FACA protein).
 GN FANCA OR FAA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 ON NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Lymphoblast;
 RX MEDLINE=97051928; PubMed=8896563;
 RA Lo Ten Foe J.R., Roosiems M.A., Bosnyan-Collins L., Alon N., Wilker M., Parker L., Lightfoot J., Carreau M., Callen D.F., Savota A., Cheng N.C., van Berkel C.G.M., Strunk M.H.P., Gille J.J.P., Pals G., Krunt F.A.E., Pronk J.C., Arwert F., Buchwald M., Jöndt H., "Expression cloning of a cDNA for the major Fancconi anaemia gene, FAA.", Nat. Genet. 14:320-323(1996).
 RL [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=97312685; PubMed=9169126;
 RA Ianzano L., D'Agostino M., Centra M., Savino M., Leyran O., Auerbach A.D., Cleton-Jansen A.M., Doggett N.A., Pronk J.C., Tipping A.J., Gibson R.A., Mathew C.G., Whitmore S.A., Apostolou S., Callen F.C., Zeltan L., Savota A.;
 RA "The genomic organization of the Fancconi anemia group A (FAA) gene.", Genomics 41:309-314(1997).
 RL [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Riche D.O., Bruce D., Mundt M., Doggett N., Munk C., Saunders E., Robinson D., Jones M., Buckingham J., Chasteen L., Thompson S., Goodwin L., Bryant J., Tesmer J., Melnick L., Longmire J., White S., Ueng S., Tatum O., Campbell C., Fawcett J., Maltbie M., Daaven L., "Sequencing of human Fancconi anemia complementation group A gene genomic region.", Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
 RL [4]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Cervix;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Struhsberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Wang J., Hsieh F.,

RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schaefer T.E.,
RA Brownstein M.J., Usdin T.B., Toohy J.K., Cantucci P., Sprengel C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez R.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzyzanski M.I., Skalska U., Smalins D.E.,
RA Scheraga A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RL [5]
RN SEQUENCE OF 491-571 AND 610-671 FROM N.A.
RX MEDLINE=98389634; PubMed=9721219;
RA Centra M., Memo E., D'Apollito M., Savino M., Ianzano L.,
RA Notarangelo A., Liu J., Doggett N.A., Zelante L., Savola A.,
RT "Fine exon-intron structure of the Fanconi anemia group A (FAA) gene
and characterization of two genomic deletions.";
RL Genomics 51:463-467(1998).
RN [6]
RX SEQUENCE OF 491-542 FROM N.A.
RA MEDLINE=98375680; PubMed=9711872;
RT Levan O., Doggett N.A., Auerbach A.D.,
RT "Identification of Alu-mediated deletions in the Fanconi anemia gene
FAA.";
RL Hum. Mutat. 12:145-152(1998).
RN [7]
RP SUBCELLULAR LOCATION, AND MUTAGENESIS.
RX MEDLINE=98414611; PubMed=9742112;
RA Naef D., Kupfer G.M., Suliman A., Lambert K., D'Andrea A.D.,
RT "Functional activity of the Fanconi anemia protein FAA requires PAC
binding and nuclear localization.";
RL Mol. Cell. Biol. 18:5952-5960(1998).
RN [8]
RP PHOSPHORYLATION.
RX MEDLINE=99007271; PubMed=9789045;
RA Yamashita T., Kupfer G.M., Naef D., Suliman A., Joenje H., Asano S.,
RA D'Andrea A.D.,
RT "The Fanconi anemia pathway requires FAA phosphorylation and FAA/PAC
nuclear accumulation.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:13085-13090(1998).
RN [9]
RP VARIANTS FA, AND VARIANTS.
RX MEDLINE=98058767; PubMed=9371798;
RA Levan O., Erlich T., Magdalena N., Gregory J.J., Batish S.D.,
RA Verlinder P.C., Auerbach A.D.,
RT "Sequence variation in the Fanconi anemia gene FAA.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:13051-13056(1997).
RN [10]
RX VARIANTS FA, PRO-1110 AND GLY-1117.
RA MEDLINE=99225041; PubMed=10210316;
RT "A patient-derived mutant form of the Fanconi anemia protein, FANCA,
is defective in nuclear accumulation.";
RL Exp. Hematol. 27:587-593(1999).
RN [11]
RX FUNCTION: DNA repair protein that may operate in a postreplication
repair or a cell cycle checkpoint function. May be involved in
interstrand DNA cross-link repair and in the maintenance of normal
chromosome stability.
RN [12]
RX SUBUNIT: Belongs to a multisubunit complex composed of FANCA,
FANCF, FANCG and FANCD proteins.
RN [13]
RX SUBCELLULAR LOCATION: Nuclear (major) and cytoplasmic (minor).
RN [14]
RX ALTERNATIVE PRODUCTS:
RX Event=Alternative splicing; Named isoforms=2;
RX Name=1;
RX IsoId=O15360-1; Sequence=Displayed;
RX Name=2;

CC IsoId=O15360-2; Sequence=VSP_007039;
CC Note=No experimental confirmation available;
CC PTM: Phosphorylation primarily on serine residues. Phosphorylation
CC is required for the formation of the nuclear complex. Not
CC phosphorylated in cells derived from groups A, B, C, E, F, G, and
CC H.
CC [15]
CC DISBASE: Defects in FANCA are the cause of one of the 8
CC complementation groups of Fanconi anemia (FA) [MIM:227650], an
CC autosomal recessive disorder characterized by progressive
CC pancytopenia, a diverse assortment of congenital malformations,
CC and a predisposition to the development of malignancies. At the
CC cellular level it is associated with hypersensitivity to DNA-
CC damaging agents, chromosomal instability (increased chromosome
CC breakage), and defective DNA repair. FANCA defects account for 65%
CC of FA.
CC [16]
CC DATABASE: NAME:Fanconi Anemia Mutation Database;
CC WWW="http://www.rochelellier.edu/fanconi/mutate/";
CC [17]
CC DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC WWW="http://www.infobiogen.fr/services/chronocancer/Genes/FA11D102.html".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X59226; CAB67610.1; -;
CC EMBL; 283067; CAB05445.1; JOINED.
CC EMBL; 283068; CAB05445.1; JOINED.
CC EMBL; 283069; CAB05445.1; JOINED.
CC EMBL; 283070; CAB05445.1; JOINED.
CC EMBL; 283071; CAB05445.1; JOINED.
CC EMBL; 283072; CAB05445.1; JOINED.
CC EMBL; 283073; CAB05445.1; JOINED.
CC EMBL; 283074; CAB05445.1; JOINED.
CC EMBL; 283075; CAB05445.1; JOINED.
CC EMBL; 283076; CAB05445.1; JOINED.
CC EMBL; 283077; CAB05445.1; JOINED.
CC EMBL; 283078; CAB05445.1; JOINED.
CC EMBL; 283079; CAB05445.1; JOINED.
CC EMBL; 283080; CAB05445.1; JOINED.
CC EMBL; 283081; CAB05445.1; JOINED.
CC EMBL; 283082; CAB05445.1; JOINED.
CC EMBL; 283083; CAB05445.1; JOINED.
CC EMBL; 283084; CAB05445.1; JOINED.
CC EMBL; 283085; CAB05445.1; JOINED.
CC EMBL; 283086; CAB05445.1; JOINED.
CC EMBL; 283087; CAB05445.1; JOINED.
CC EMBL; 283088; CAB05445.1; JOINED.
CC EMBL; 283089; CAB05445.1; JOINED.
CC EMBL; 283090; CAB05445.1; JOINED.
CC EMBL; 283091; CAB05445.1; JOINED.
CC EMBL; 283092; CAB05445.1; JOINED.
CC EMBL; 283093; CAB05445.1; JOINED.
CC EMBL; 283094; CAB05445.1; JOINED.
CC EMBL; 283095; CAB05445.1; JOINED.
CC EMBL; 283151; CAB05445.1; JOINED.
CC EMBL; AC005360; AAC28751.1; -;
CC EMBL; AC005565; AAC33304.1; -;
CC EMBL; AC005567; AAH03401.1; -;
CC EMBL; BC008879; AAH08979.1; -;
CC EMBL; AF0225084; CAI12393.1; -;
CC EMBL; AF0225085; CAI12394.1; -;
CC EMBL; AF054569; AAC28331.1; -;
CC Genew; HGNC:3582; FANCA.
CC MIM; 607139; -;
CC MIM; 227650; -;
CC GO; GO:0005737; C:cytoplasm; TAS.
CC GO; GO:0005634; C:nucleus; TAS.
CC GO; GO:0006281; P:DNA repair; TAS.
CC GO; GO:0006461; P:protein complex assembly; TAS.

DR InterPro: IPR003516; Fanconia.
DR Pfam; PF03511; Fanconia; 1.
DR PRINTS; PR00826; FANCONIAGENE.
KW DNA repair; Nuclear protein; Phosphorylation; Disease mutation;
KW Polymorphism; Alternative splicing.

Query Match 1.8%; Score 138; DB 1; Length 1455;
Best Local Similarity 19.3%; Pred. No. 0.93; Mismatches 225; Conservative 139; Indels 408; Gaps 58

Qy	143	IOAQLERNNKENIQLDROFDMKLIINFFSNASQCYLCPSAQONLAAQGDLMG----	198
Db	366	LSAEELVGHIOLEVLETO-EVHMQRSLSFVSLVCFE-PEAQO-LLEDWARLMAQAPE	420
Qy	199	-----VLDYLOQEVPRAFVNLVLDSEVAEVS-ROYNHGTWLSAPAEPCSCSEET	245
Db	421	SCOLDMSVTAFLVVRQALBESPSAFLSADMFKASFGSTRGYNH-----CSKA	469
Qy	246	-----TLAKYVMQMSYOEAMNSLLASSRYSEQBSFTVVPQFFYET--TESLSEDERL	298
Db	470	LVEFLFTLSLELVP-----FESPRYIQVHILHPLVPSKRSRLTDYISLAKTRL	518
Qy	299	QDSTTLMLHMN-----RMERPAGEKDEPLS-----VGH-GRPMKCPGSESYL	341
Db	519	AD---LKVSYIENMGLYEDLSAGDITTEHNSQALQDVEKAIWFEHTGNIPTYVMEASIFR	575
Qy	342	FSYKNSNYLTRLQKPODKLEVBEGAEIRCPDKDSDPYTPTSVHRL-----KPADI	391
Db	576	RPYVSHFLPALLTRPVLPKVPDSRVAFIEBKLKADKIPRELSYGYCAGASAEKPEPD	634
Qy	392	NVIGALDLSLTAGAGASTPENVLDVLTYQKGLSWSVGDENIGTVTTLANIIRBFNPSL	451
Db	635	-----AALGVRAEPNS-----AEEPQLQTLAALGELR-----	661
Qy	452	KGFVGTGKETSPN-----AFINQ---AVAGRAEDLPYQARRLVYDMKNDTRIHF	499
Db	662	-----ASMTDPSQRDYISAQVAVISRLRAVLGHNNEDSSVEISK-IQLSINTPRLEP	713
Qy	500	QEDWKIITLPIGNDLCDPNDLHVHSPQNFNDIGKALDILHAEPRAFNVLVTLEIV	559
Db	714	REHNAVLDL-----LTSFCQNLMAAS-----SVAP-----	738
Qy	560	NLABLYOEKVVYCPRMILRLSLC---PCVLKFDNDSTELATLIEFNKKFOEKTQOL	611
Db	739	-----PERQPMALFLVRTWCGHVLPAVL-----TRLQQLRHQPSLSAPHYLGAA	786
Qy	612	--IESGYDTRREDTIVVQPFPEVNDMPKTSBGLP-----DNSFPAPOCF	654
Db	787	LAVHLG--ESRSALPEV-----DVGRPARGAPLVRPALFDSLLTCRTTRDSLFF--CLK	835
Qy	655	-----HFSSKSHRAASALMMNMLPEVQOKTRHFKFNKINTCPNOVQOFLRT	703
Db	836	FCTAALISYLSLCKPSSQSRDITCGLSLPOLIK-----KFOFLMPLPBEAQOPLSEE	886
Qy	704	YKSMQCHGTWLPGRDRAPSAL---HPTSVHALRPADIQVVAALGDSLTAANGIGSKPD	759
Db	887	DVASLWRPRLHPBADQORALSLMTHTRTFEVLEKEEDVHL--TYQDWLHLELEIOPRAD	944
Qy	760	DLPRVTQO-----YKGSYSAGGDSLENV--TTLPNILIREPRNRLTGAVGCD	807
Db	945	ALSTERQDPRHQWAIHFNHLPRESSAGGCGDGLDAACTIVNALAMDFFGSSRSY---D	999
Qy	808	ANDTNAFLNQAHPKAKA--EDLMSQVOTLMQMKMDKRVNFEHEDMKVITVLIGSDSLCDY	865
Db	1000	HSEN-----SOLVFGRTGNEIDIISRLQEMVADLE-----LQOB---LIYELGHTP	1042
Qy	866	CTBSNLSAANFV-DHLRNALDVL-----HREV---PRVLVNLVDPLNPTMR	909
Db	1043	-----SOEHFLFEIRFRRLQALTSQMSVAAISQORRELLMYKRIILRL-----	1085
Qy	910	QVFLGNPDKCPVQOQASVLCANCVLLIRENSQGLAELEAFSRAYSRSMELVSGSKYDPOED	969
Db	1086	-----PSSVLCGS--SFQABQITPACBQFPHLVNSBMRKFCSHGALTYQD--	1129

```

QY 970 FSVLQGFPPNIOLEPVLADLPD--TSFPADCIHPNOQKHSOLARL--WTNNLEPAGSK 1026
Dh 1130 ----ITAFIFPGLNACLRSDRPSLMDVFIKAC--QTKCPILITSLVWMPSPLEPV--- 1180
QY 1027 TETLLDRAEMPICTPQONEBPFLTRPNSNTYPIKPAIE--NMGSDFLCTE----- 1075
Dh 1181 -----LTCWRHRHCQSPLEPRE-----LQKQEGRQFASDPLSPESAASPAIPND 1223
QY 1076 WKASNSVPTSVHQLRPADIK-----VVAALGDSLTITVAGARPNN 1114
Dh 1224 WLSAAALHFAPIQOVRREENIRKQKJDCEREELLVLPFPSLGLGLSSHLTS-----NS 1277
QY 1115 SSDLPPTSM-----RGLSW 1127
Dh 1278 TTDLPRKAFHVCAAILBECLERKRLSW 1302

```

RESULT 4	GLI4_XENMLA	STANDARD;	PRT; 1361 AA.
ID	GLI4_XENMLA		
AC	Q91661;		
DT	01-NOV-1997 (Rel. 35, Created)		
DT	01-NOV-1997 (Rel. 35, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
DE	zinc finger protein GLI4 (Neural specific DNA binding protein XGLI4) (XGLI-4).		
CN	GLI4.		
OC	Xenopus laevis (African clawed frog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;		
OC	Xenopodinae; Xenopus.		
OX	NCBI_TaxId=8355;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	MEDLINE=97346726; PubMed=9203143;		
RA	Marine J.C., Bellefroid E.J., Pendeville H., Martial J.A., Pieler T.;		
RT	"A role for Xenopus Gli-type zinc finger proteins in the early embryonic patterning of mesoderm and neuroectoderm.";		
RL	Mech. Dev. 63:211-225(1997).		
CC	-1- FUNCTION: HAS AN ESSENTIAL ROLE IN THE EARLY EMBRYONIC PATTERNING OF MESODERM AND NEUROCTODERM.		
CC	-1- SUBCELLULAR LOCATION: Nuclear (Potential).		
CC	-1- SIMILARITY: BELONGS TO THE GLI FAMILY OF C2H2-TYPE ZINC-FINGER PROTEINS.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-eb.ch/announce/ or send an email to license@isb-eb.ch).		
CC	-----		
DR	EMBL; U42462; AAA98467.1; -		
DR	PIR; T30884; T30884.		
DR	HSSP; P08151; ZGLI.		
DR	InterPro; IPR007087; Znf_C2H2.		
DR	Pfam; PF00096; ZF-C2H2; 5.		
DR	SMART; SM00355; Znf_C2H2; 5.		
DR	PROSITE; PS00028; ZINC_FINGER_C2H2_1; 4.		
DR	PROSITE; PS0157; ZINC_FINGER_C2H2_2; 5.		
KW	Zinc-finger; Metal-binding; DNA-binding; Transcription regulation;		
KT	Nuclear protein. Repeat.		
FT	ZN FING 289 314 C2H2-TYPE.		
FT	ZN FING 322 349 C2H2-TYPE.		
FT	ZN FING 355 379 C2H2-TYPE.		
FT	ZN FING 385 410 C2H2-TYPE.		
FT	ZN FING 416 441 C2H2-TYPE.		
SQ	SEQUENCE 1361 AA; 149554 MW; 03AC399BFA4CC40C CRC64;		
Query Match	1.7%; Score 131.5; DB 1; Length 1361;		
Best Local Similarity	19.7%; Pred. No. 2.3;		

Matches 249; Conservative 140; Mismatches 498; Indels 379; Gaps 58;

QY 56 NMPSKVS-----LKPSDIK-----VAIGNLEIPRPGTGDELEKOWTERPOOV-----103

Db 10 NSPTLMSISARGISPEVAHEHLKERIGYGLAPPPPGT-----TPREYCHQMA 59

QY 104 -----MGWTVLS-----DIIRFSSPVMPVCHTKRVIPHDGA 139

Db 60 FLASHSPYELLVQSAAGNTSHLDHYLTPMDVSRSSPRVTPRL--SKKALSTISPLS 117

QY 140 DLMTIQAEVLVRNKENIQLDFQDKILINFPNSAQCYLCPSAQONGLAAGVD-----194

Db 118 DASIDQLMTIRTSPTNSL-----YAYINSSSSASGSYGHLISAGASIPASFS 165

QY 195 -----ELMGVLVLYQGEVPRAPVNLVDLSVAEVSQYHG--TWLSRPEPCNCSFTT 246

Db 166 PIHKPGLSALSO-QSLSSSFHTPLHPSPTFASRQOQALSNAPAPPSSNSAPDS 224

QY 247 RLAKVVMQMSYOEAMNSL--ASSRYSSESFTVVF-QP-----282

Db 225 VLAKVSESASVSTVQVHHRKSVKTEBEADSVRFQPDHLDLDEKDECKQRE 284

QY 283 FFEYETPSLHSEDPRLQDSTTLAMHLMNRMPEAGEKE-----PLSVKH--327

Db 285 HIYETNCHMDGCKSEFPTQDQLVHINNDHIH--GEKKEFCWQDCSRQCKPFAQYML 342

QY 328 -----GRPKKPSQ-----ESPILFSRNSN-----YITRL 353

Db 343 VHMRRHTGKPKHCTFEGCFKAYSRLNKLTHLSHTGKPYVCHDEGCKAFSNASDR 402

QY 354 QKPDCKLEVEGAEL--RCPDKPSDVTPTSV-----HRLKPADINVGALGDS 400

Db 403 AKQGNRHSHNEKPYICVPCCTKRYTDPSSLRKVKTVHQBPAHVTKKHNDIIGK--PS 460

QY 401 LTAGNGAGSTPGVNLVLTQYRGLSMSVGDE-----NIGVTTLANTLREFPNLSKGSV 456

Db 461 LPRKNG-----DNEASAKLSGREHSDSVRQEHCLQTRITKTEDNMQSSP--GGQS 512

QY 457 GTGKERSPNALNQAVG-----GRADLPVQARLVLDLKN-----D 494

Db 513 SCSEBSPYGNNTNIDSGVDVSLAMQSLDLFGLLETSPTVDSTVSWORSGRPATPT 572

QY 495 TRIHFOE-----DMKITLFLIGNDLDFCNLDLVHSPONFNDNIGKALDILAE 544

Db 573 QRHSAETGAEREIENDNERFLIYPRNATQ--NTRLPLISANGDVIGVESVAIN--629

QY 545 VPRAF--VNLVTVLEIVNRELQEKVYCPRMILRSLCPVLCFDDNSTELATLIEFN 601

Db 630 -PRAIELSNMDVTMMQNLNERR-----DSTSSTL-----657

QY 602 KKQOEKTHQLESGRYTREDFTYVVOFPFENVDMKTSF--GLPDSFPAPOCFHRSK 659

Db 658 -----SSAYTSRR--SSGISPYFSRRSSSTSOFGGLNNSSSADSYPISTD 703

QY 660 SHSRAGALNMNMLEPVGQKTRHKPFENKINIC-----PNOVOPFLRTYKSMQ 709

Db 704 ASRSSSSASGSGCLPNLINTLTPAOHYLKAKYAAATGCPPTPLPNDRIGLRKLSLMD 763

QY 710 GHGTWLP-----CRDRAPSA-LHPTSVALRPAIDQVVAALGDSITLGAAGISKP 758

Db 764 GADPPLPFPRLQVPRRCDSGGAGLTPMYPHI--PGNNSRRAS--DVARRTAGIDKP 820

QY 759 DDLPDVTYQRGSLYSAGDSLENVTL--PILRFPRNLITGAVGTGDAN-----809

Db 821 --LPRFSRFH-----SNMSNMTLHPPSLSEBRNGQLQHYTSDGGLHRYVSPR 867

QY 810 -----DTAFL-----NOAVFGAKAEDLM-----SOVQTLMOQMKDHRVNFHEDMVIIVLI 857

Db 868 PPSISEVAMBAISCDAVREGD--DDMLPDDVVQYIRSQRRAPBQULOTERS-----920

QY 858 GGSGLCDYCTDSNLSAA-----NFVDHLRNALDVLHREVPV-----LVNLVDPLNPT 906

Db 921 -----SPARNLQSNVTKSFHNNTPPQPRAPGAYLERNFPALAECCGGATANNQDNMPV 972

QY 907 IMQVFLGND--KCPVOASVLCNCLTLRENSQELARLEAPSRAYRSMRELVSGRY 964

Db 973 QMNEVSSGVTVDSLDLEKQQAFA---GNLAVQOKQFQAYQSNQAPMPQRAHIMQGOE 1029

QY 965 DTQEDSPVYLQPF-----FQNIQLPYLADGLDPTSEFAPDCI--HPNQFHSOLA 1012

Db 1030 SVQRIINISVQGRFNVLQORQOQMSQCI--VSSDFIQORYVSQSQMSLSRAMQEGOSQIS 1088

QY 1013 RALMTMLBPLGSKT-----ETLDDLAEMPIPTPTO-----NEPFLRTPNSNTYP 1059

Db 1089 PSC-NNMVERPGVHTHAPSNTLHQRLAVHAPGTGFANNFSVNODGLHP--NAYTVOP 1146

QY 1060 IKPAIE 1065

Db 1147 QKNGLE 1152

RESULT 5

C27A_MOUSE STANDARD; PRT: 1155 AA.

AC P79621; 046787; 078036; 078109; Q31115; Q9TPP1;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-SEP-2003 (Rel. 42, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE MHC class II transactivator (CITRA).

GN MHC2TA OR CITRA OR C27A.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

XP [1]

RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3), TISSUE SPECIFICITY, AND INDUCTION.

RC STRAIN=BALB/c;

RA MEDLINE=97327562; PubMed=9184229;

RT "Mhlethaler-Motter A., Otten L.A., Steinle V., Mach B.; Expression of MHC class II molecules in different cellular and functional compartments is controlled by differential usage of multiple promoters of the transactivator CITRA.";

RL EMBO J. 16:2851-2860 (1997).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM 2).

RC STRAIN=NOD; TISSUE=Spleen;

RX MEDLINE=97152412; PubMed=8995190;

RA Stims T.N., Elliott J.E., Ramasser V., Denney D.W. Jr., Halloran P.F.; "Mouse class II transactivator: cDNA sequence and amino acid comparison with the human class II transactivator.";

RT Immunogenetics 45:220-222 (1997).

RN [3]

RP SEQUENCE OF 955-1097 FROM N.A.

RC STRAIN=BALB/c; TISSUE=Thymus;

RX MEDLINE=9618886; PubMed=8620527;

RA Panelli M.C., Wang E., Shen S., Schluter S.F., Bernstein R.M., Harsh E.M., Stoeck A., Gangavalli R., Barber J., Jolly D., Apoptocyte E.T.;

RA "Interferon gamma (IFNgamma) gene transfer of an EMT6 tumor that is poorly responsive to IFNgamma stimulation: Increase in tumor immunogenicity is accompanied by induction of a mouse class II transactivator and class II MHC.";

RT Cancer Immunol. Immunother. 42:99-107 (1996).

RL -FUNCTION: ESSENTIAL FOR TRANSCRIPTIONAL ACTIVITY OF THE HLA CLASS II PROMOTER; ACTIVATION IS VIA THE PROXIMAL PROMOTER. NO DNA BINDING OF IN VITRO TRANSLATED CITRA WAS DETECTED. MAY ACT IN A COACTIVATOR-LIKE FASHION THROUGH PROTEIN-PROTEIN INTERACTIONS BY CONTACTING FACTORS BINDING TO THE PROXIMAL MHC CLASS II PROMOTER, TO ELEMENTS OF THE TRANSCRIPTION MACHINERY, OR BOTH. ALTERNATIVELY IT MAY ACTIVATE HLA CLASS II TRANSCRIPTION BY MODIFYING PROTEINS THAT BIND TO THE MHC CLASS II PROMOTER (BY SIMILARITY).

CC -SUBCELLULAR LOCATION: Nuclear.

CC -FUNCTIONAL PRODUCTS:

CC -Event=Alternative splicing; Named isoforms=3;

CC Name=1; Synonyms=1;

CC		IsoId=p79621-1; Sequence=Displayed;
CC	Name=2; Synonyms=III;	
CC	IsoId=p79621-2; Sequence=VSP_007214, VSP_007215;	
CC	Name=3; Synonyms=IV;	
CC	IsoId=p79621-3; Sequence=VSP_007216;	
CC	-1-TISSUE SPECIFICITY: Isoform 1 is expressed at very high levels in dendritic cells, at very low levels in spleen and thymus and is not detected in other tissues. Isoform 2 is detected at high levels in spleen and tonsil as well as in a number of B-lymphocyte cell lines, and at very low levels in dendritic cells.	
CC	-1-INDUCTION: By interferon-gamma.	
CC	-1-SIMILARITY: Contains 5 leucine-rich (LRR) repeats.	
CC	-1-SIMILARITY: Contains 1 NACHT domain.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as it remains free of charge. The database modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@ib-eb.ch).	
CC	-----	
DR	EMBL; AF000006; AAB92364.2; -	
DR	EMBL; AF000007; AAB92365.1; -	
DR	EMBL; AF042158; AAC34367.1; -	
DR	EMBL; AF042159; AAC34367.1; -	
DR	EMBL; AF100709; AAP06838.1; -	
DR	EMBL; AF100710; AAP06839.1; -	
DR	EMBL; U06053; ABA48859.1; -	
DR	EMBL; U46562; AAB05004.1; -	
DR	MGI; MGI:108445; C2ca.	
DR	InterPro; IPR001611; LRR_RNinh sub.	
DR	InterPro; IPR003590; LRR_RNinh sub.	
DR	InterPro; IPR007111; NACHT_NTRase.	
DR	Pfam; PF00560; LRR; 1.	
DR	SMART; SM00368; LRR_RI; 4.	
DR	PROSITE; PS50837; NACHT; 1.	
KW	Transcription regulation; Activator; Nuclear protein; ATP-binding; Leucine-rich repeat; Repeat; Alternative splicing.	
KM	REPEAT 175 198 LRR 1.	
FT	REPEAT 243 269 LRR 2.	
FT	DOMAIN 439 749 NACHT.	
FT	REPEAT 979 1000 LRR 3.	
FT	REPEAT 1008 1031 LRR 4.	
FT	REPEAT 1039 1062 LRR 5.	
FT	DOMAIN 129 215 ASP/GLU-RICH (ACIDIC).	
FT	NP_BIND 445 452 Missing (in isoform 2).	
FT	VARSPLIC 1 77 /FTid=VSP_007214.	
FT	SLOAFTERGTSRDHG -> MRCLVPSSGSYLEPQ (in isoform 2).	
FT	VARSPLIC 78 94 /FTid=VSP_007215.	
FT	VARSPUBLIC 1 101 Missing (in isoform 3).	
FT	CONFLICT 994 994 /FTid=VSP_007216.	
FT	SEQUENCE 1155 AA; 127528 MW; F3FPO5DEBFBFCR71 CRC64;	
SO		
Query March	1.6%; Score 128; DB 1; Length 1155;	
Best local similarity	22.9%; Pred. No. 3;	
Matches 102; Conservative	47; Mismatches 153; Indels 144; Gaps	
OY	39 ETLNKSPPCPNP-KLAGVNMP-----SKSVSLKPSDKIFAALGNLEIPDP-----85	
Db	337 ETEPTSPSCQGEPPSSIKLKPKWEPAVRPFQHSIQD---KKYA-----LPQSPRGPLVA 387	
OY	86 -----GTG-----DLEKQMTPEPQQVCMG-----VMTVLSDITIRYSPSYVMPC 126	
Db	388 VELVVARLERGSSNKQERELATPWTER--QLAHGLAEVLQVNSDCRR----PGETQYVA 442	
OY	127 HTGK-----RVTPHDGAEDLIQAQELVRNNKENJOLDFOGDMLINVF-----171	
Db	443 VLGRAGGCKSMWARTVSH-----TWAGQ-----LLQYDF-----VFVYPCHCL 481	

```

OY 172 SMAAGC-----LCPAQQNGLAAGVDELMGVDYLQGEVPAFVNLVDLSEVAVSRQ 226
Db 482 DRPGDTYHLRDLCLPPSLQ-PLAND--DE---VDYIVRQPDRLVLLIDAFEELEAGCL 535
OY 227 YHGTWLSPPAPPCPCSEBETTRALAVVWQMSYQEAWNSSILASRYSEQSEFT---VFQPF 283
Db 536 LHGCGGSLSPRCGL-----RGLLACIPIKRLRGCTLLLTARPRGLAQSLSKADAFIE-- 590
OY 284 FYETTPSLHSEDPRLQDSTLTAHLMLNRMWEPAGEKDEPLSVKGRPKKCPSGSEPLYFS 343
Db 591 ----VPSFSTK---QAKTVRRHYFEN--SGTAGNQDKALGLLEQPLLCC-----SS 631
OY 344 YRNSNYLIRLQKQPKDKLEVRGCAIEICRQDKDPSDTPVRSVHRLKPADINVGALGDSLTA 403
Db 632 YSHSPVPCRAVCQSLSKALLBEGTAPQLP-----CTLGLTVSLIG 671
OY 404 GNGAGSTPGNVLDTQYRGLSWAG 429
Db 672 PAQNSPPG-----ALVELAKLAWELG 693

RESULT 6
YFAS_ECOLI
ID YFAS_ECOLI STANDARD; PRT; 1534 AA.
AC P76464. P76465.
DT 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein yfas precursor.
GN YFAS OR B2227/B2228.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
NCBI_TaxID=562;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655.
RX MEDLINE=9742617; PubMed=9278503;
RA Blatcher F.R., Plunkett G. II, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collins-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Alpa H., Baba T., Hayaashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nishimoto H.,
RA Oshima T., Oyama S., Saito N., Sampel G., Satoh Y., Sivasundaram S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horiuchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT K12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features.";
RL DNA Res. 4:91-113(1997).
[3]
RP CONCEPTUAL TRANSLATION.
RA Coudert E.;
RL Unpublished observations (MAR-2002).
CC -1- SIMILARITY: BELONGS TO THE UPF0192 FAMILY.
CC POSITION 1368 WAS TRANSLATED AS TRP TO RESTORE THE SIMILARITY WITH
CC THE C-TERMINAL REGION OF OTHER HOMOLOGS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcement/)

```

or send an email to license@slb.ch.

CC -----
 DR EMBL: AEO00312; AAC75288.1; ALT INIT.
 DR EMBL: AEO00312; AAC75287.1; ALT INIT.
 DR EMBL: D90854; -; NOT_ANNOTATED_CDS.
 DR EMBL: D90851; -; NOT_ANNOTATED_CDS.
 DR EcoGene: EG14081; yfas.
 KM Hypothetical protein; signal; Complete proteome.
 FT SIGNAL 1 38 POTENTIAL.
 FT CHAIN 39 1534 HYPOTHETICAL PROTEIN YFAS.
 SQ SEQUENCE 1534 AA; 169845 MW; 73B71431C1F8B6 CRC64;

Query Match 1.6%; Score 126.5; DB 1; Length 1534;
 Best Local Similarity 18.8%; Pred. No. 5.9;
 Matches 278; Conservative 174; Mismatches 556; Indels 473; Gaps 68;

QY 161 QFHWKLINFFSNASQCYLCPSAQOQGLA-----AGVDELMGVLDYIOQE 206
 DB 9 QFHWHL-SPRFSGAIAACLSLIVGTGLANNADSLPSSNAPAPAGCTFFLIADSSFSSE 67
 QY 207 VPRAFVVLVDSEVAEVSROYH-----GTWLSRPPPCSCSEETTRIAKVMQWQY-- 257
 DB 68 EAK-----VLEBAPGRDYRRIOHEEYGGVDVRLYRIPMAFLRQOKLHRTIVQPOLG 122
 QY 258 -----QEAANSLASSRYSEBSFTVVPFPFFETTPSLHSIDPRLQDSTTLAMHLMN 310
 DB 123 DGLNLTLTWLMDMWYKGRSRWQRTFSSQSKQNTQALPELQGNALIIKPSR-----YVQN 178
 QY 311 RMNEPAGEKDEPLSVKGRPMKCSQSPYLFSYRNSNYLTROKLPQDKLEVRGAEIRIC 370
 DB 179 NQSPPL-KKYPVAKQPRYPV--WQAKP-----EPQGVVL-EGA--- 214
 QY 371 PDDPSPTVTSVRLRPADINIVIGALGDSLTAGNGAGSTPPGNVLDVLTQYRGLSWSVGG 430
 DB 215 -----SSNFISSPOGNIYIFLQ----- 232
 QY 431 DENIGVTTLANILREPNPSLKPSVGTGKETSPPAFLQAVAGRAEDLPVQARRLVDL 490
 DB 233 -----QERGVLVEAMVG-----YRATTVV 253
 QY 491 MKQDTRHNPEDWKITLFTIGNDLCD-----FCNDLVHYSQONTNIGKALDILHA 543
 DB 254 FVSDTVLALSKVSGKELVMTAGKKQGAKEPSEILMTDGLVMTRGTDDSG--TLQQLHI 312
 QY 544 EVPRAPV-----NLVTVLEIVNLRELIOEKKVCPRMILRLSLCPVLKPEDNSTELATL 597
 DB 313 SPEKSYTLGDAGSGVFS-----NFFYESEITNTLYLFTDRPLRYAGRDVVKV-- 364
 QY 598 IEFNKKFOEKTHQ-----LIESGRVDTREDFTVVQPFENVDMPTSGE--LPDNSFFAP 651
 DB 365 --IGREHDPDHSPIVSAPAKSLVDANGSLLOTVAVTLDARRGGGSPFLPENAVAGG 422
 QY 652 DCFHSSKSHRPAASALMNMLEP--VGOKTRHKTE-----NKINITCPN----- 695
 DB 423 YELRLAARNQVSSSFFRVANYIKHFEIGLALAKKEPKTGAVSGKQLDLYPDEPYVNA 482
 QY 696 QVOPFLATYNSMGCHTWTLPDRAPASALHPTSVHMLRPADIQVVALDGLSLTAGGIG 755
 DB 483 RVQSLSLAQQLSMVGN-----DLRYAGRPVS-----LBSSETVSPASG 521
 QY 756 SKPDDLDVTVYQGLSYSGAGDGLSNVTLPLNILEFNNLTGAVGT-----GDANDT 811
 DB 522 HVALNLPRAAKPSRYLTLVSASDGAAYRTTTRKILIE--RGLAHYGLSTAAQVNSGES 579
 QY 812 NAFINQAVPGAK-----ADLMSQVOTLMQKMDHRVNF-----HEDW 850
 DB 580 VFRFYALLESSKQVPVTEWMLRLEDRSHSGELSPSGK--SETVNFAKRGVNLTLRKDG 638
 QY 851 KVTIVL---IG-----SDLCYCTDSNLYSANFPVDHLRNALDVLHREPRVLVNLVD 901
 DB 639 LILAGLSHAIVSGKSTAHGTVDIVAAKTLVQPG-----ETAKMLITFPE 683
 QY 902 FLNPTIRQVFLGNPKCPVQQAIVLCNVCV--LTLR--ENGQILARLEAFSRAVRSWMRE 957

DB 684 PIDEALL-----TLERDRVEQOOSLSHPANWTLQRLANDTOYEARVPV----- 726
 QY 958 LVNSGRDIOEDSFVYL-----QPFQNIQCPVLADGLPDTSFPAFDIHNQKHSQAR 1013
 DB 727 ---SNSFAPNITSVLTTRNGQVSPNAGIKVAVPOL-----DIRVTKDTHYOPGE 775
 QY 1014 ALMTNMLEPISGTEETDLR-----AEMPTCTQNEPFLRTPRNSNYT----- 1057
 DB 776 LVNVELTSLKGRVSAQLVGVVDEMIYALQELIANICKFFYPLGRNNVRRSSLSFI 835
 QY 1058 -----YPIKAIEHWGSDFLCTEWMKASVPTSVHOL--RPADIKV--VAALGDSLTTA 1107
 DB 836 SYDQALSSSEVADG-----ATNRSEBRVQWLEPRRREEVDTAAWMPSLTTD 881
 QY 1108 VGARPNNSDLP---TSWRGLSWSIGDG-----NLEHTTTLPNILKKNPY 1151
 DB 882 KQGAFTFLMPSLTRWRTTARGMNGDGLVGGRAVLRSEKLYMKMSPTVY----- 935
 QY 1152 LLGFSTWEGTAGLVNAAGARAPDPAQMDLVERMKSPDINLEKDWKLYTLFTIGVN 1211
 DB 936 -----RVGDKRPAAGLPIFSQDNEP-----VALYTKFAG-- 964
 QY 1212 DLCHYCENPRAHLATEYV---QHIOALDILSEELP-----AFV-NVEVME 1255
 DB 965 --AEMRQTLTLHKGANYSILTQNIQOS--GLLSAELIQNGQVODSITKLSFVDSNMPVEQ 1021
 QY 1256 LASLYQGGGKCMALQAQNNCTCLRHQSLSLEQELKVN-----W-----NLQHCIS 1303
 DB 1022 QKNVMLGSGDVALMLPEQAQSNIRL---QSSETPOELFRNNLDALVDEPVGAVINTGRLI 1078
 QY 1304 SFESY-W-----HOYQREDAVVVQPFQNTLTPLENRG--DTDLTFSESDCFHSDGHA 1356
 DB 1079 PLSIAPRSLADHQAANDIRQMIQ---DNRLMLQLAGGARFTWNGED-----GNG 1128
 QY 1357 EMALAMNNMLEPVGRKTTNNFTSHRAKLCSPSPSYLYTLRNSFLPDQAEAD-- 1413
 DB 1129 DAFILAW-----AMYADWQAQSIGVTQOPR--YQGHMIDS--VAEQADNMMLH 1174
 QY 1414 --EVLVYA-----VPVAGVGLVVGIICTVYWRGRGGRRED 1448
 DB 1175 RALVLAWAQEMNLPCKT--LNGGLDEAI--ARRGTED 1209

RESULT 7
 SNE2_HUMAN STANDARD; PRT; 6885 AA.
 AC Q8WXH0; Q8NIS3; Q8NF49; Q8TER7; Q8MWM3; Q8MWM4; Q8MWM5; Q8WXH1;
 AC Q8WU50; Q9UFQ4; Q9Y2L4; Q9Y4R1;
 DT 15-SEP-2003 (Rel. 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Neprtin 2 (Nuclear envelope spectrin repeat protein 2) (Syn-2)
 DE (Synaptic nuclear envelope protein 2) (Nucleus and actin connecting
 DE element protein) (NUANCE protein).
 GN SNE2 OR NUA OR KIA1011.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 ON NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 9), FUNCTION, CHARACTERIZATION, AND
 RP INTERACTION WITH F-ACTIN.
 RX MEDLINE=22113122; Pubmed=12118075;
 RA Zhen Y.-Y., Libotte T., Munch M., Noegel A.A., Korenbaum E.;
 RT "NUANCE, a giant protein connecting the nucleus and actin
 RT cyoskeleton.";
 RL J. Cell Sci. 115:3207-3222 (2002).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 4; 5 AND 7).
 RX MEDLINE=21652858; Pubmed=11792814;
 RA Zhang O., Skepper J.N., Yang F., Davies J.D., Hegyi L., Roberts R.G.,
 RA Weissberg P.L., Ellis J.A., Shanahan C.M.;

RT "Nesprins: a novel family of spectrin-repeat-containing proteins that
 RT localize to the nuclear membrane in multiple tissues.";
 RL J. Cell Sci. 114:4485-4498(2001).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=2236983; PubMed=12408964;
 RA Zhang Q., Ragnauth C., Greener M.J., Shanahan C.M., Roberts R.G.;
 RT "The nesprins are giant actin-binding proteins, orthologues to
 RT Drosophila melanogaster muscle MSP-300.";
 RL Genomics 80:473-481(2002).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 6).
 RC TISSUE=Testis;
 RA Foustka A., Klein M., Mewes H.-W., Gassenhuber J., Wiemann S.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RX PubMed=12508121;
 RA Heilig R., Eckenberg R., Petit J.-L., Fonknechten N., Da Silva C.,
 RA Cattelico L., Levy M., Barbe V., De Berardinis V., Ureta-Vidal A.,
 RA Pelletier E., Vico V., Antougar V., Rowen L., Madan A., Qin S.,
 RA Sun H., Du H., Pepin K., Artiguenave F., Robert C., Cruaud C.,
 RA Bruns T., Jallion O., Friedlander L., Samson G., Broctier P.,
 RA Cure S., Segurens B., Aniere F., Samain S., Crespeau H., Abbaï N.,
 RA Alach N., Boscus D., Dikhoïff R., Dors M., Dubois I., Friedman S.,
 RA Gouyenoux M., James R., Madan A., Maïre-Estrada B., Mangenot S.,
 RA Martins N., Menard M., Oztas S., Ratcliffe A., Shaffer T., Trask B.,
 RA Vacherie B., Bellemere C., Belser C., Bernard-Gonet M.,
 RA Bartol-Mavel D., Boutard M., Briez-Silla S., Combette S.,
 RA Dufosse-Laurent V., Ferron C., Lechaplais C., Louesse C., Musielat D.,
 RA Magdelanet G., Pateau E., Petit E., Silvain-Trukhietz P., Trybou A.,
 RA Vega-Czarny N., Bataille E., Bluet E., Borelatis I., Dubois P.,
 RA Dumont C., Guerin T., Haffray S., Hamadi R., Munga J., Pallouin V.,
 RA Robert D., Wunderle E., Gaugier G., Roy A., Saine-Marthe L.,
 RA Verdier J., Verdier-Discala C., Hillier L., Fulton L., McPherson J.,
 RA Matsuda F., Wilson R., Scarpelli C., Gyapay G., Wincker P., Saurin W.,
 RA Queller F., Waterston R., Hood L., Weissbach J.;
 RT "The DNA sequence and analysis of human chromosome 14.";
 RL Nature 421:601-607(2003).
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM 8).
 RC TISSUE=Brain;
 RX MEDLINE=2338257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hale F.,
 RA Datchenko L., Marusik K., Farmer A.A., Rubin A.G., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo A.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loughran N.A., Peters G.J., Abramson R.D., Mullen S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchan J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [7]
 RP SEQUENCE FROM N.A. (ISOFORM 3), AND SEQUENCE OF 1-956 AND 5133-6885
 RP FROM N.A.
 RC TISSUE=Spleen, and Tongue;
 RA Jikva H., Takano J., Nomura N., Kikuno R., Nagase T., Ohara O.,
 RA Ninomiya K., Wagauma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
 RA Furuya T., Takahashi M., Kikawa E., Omura Y., Abe K., Kamihara K.,
 RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
 RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,

RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
 RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
 RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE OF 5709-6885 FROM N.A. (ISOFORM 2).
 RC TISSUE=Brain;
 RX MEDLINE=99245063; PubMed=10231032;
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosewa M.,
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIII.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 6:63-70(1999).
 RN [9]
 RP REVISIONS.
 RC TISSUE=Brain;
 RX MEDLINE=22158633; PubMed=12168954;
 RA Nakajima D., Okazaki N., Yamakawa H., Kikuno R., Ohara O., Nagase T.;
 RT "Construction of expression-ready cDNA clones for KIAA genes: manual
 RT curation of 330 KIAA cDNA clones.";
 RL DNA Res. 9:99-106(2002).
 RN [10]
 RP SEQUENCE OF 5754-6885 FROM N.A.
 RX MEDLINE=21154917; PubMed=11230166;
 RA Wiemann S., Well B., Wellenreuther R., Gassenhuber J., Glaesl S.,
 RA Anisorgo W., Boecher M., Bloeker H., Bautebachs S., Blum H.,
 RA Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,
 RA Mewes H.-W., Oestermaier B., Obermaier B., Tampe J., Heubner D.,
 RA Mambrot R., Korn B., Klein M., Foustka A.;
 RT "Towards a catalog of human genes and proteins: sequencing and
 RT analysis of 500 novel complete protein coding human cDNAs.";
 RL Genome Res. 11:422-435(2001).
 CC -1- FUNCTION: Involved in the maintenance of nuclear organization and
 CC structural integrity. Probable anchoring protein which tethers the
 CC nucleus to the cytoskeleton. Connects nuclei to the cytoskeleton
 CC by interacting with the nuclear envelope and with F-actin in the
 CC cytoplasm.
 CC -1- SUBUNIT: Interacts with F-actin via its N-terminal domain.
 CC -1- SUBCELLULAR LOCATION: Type IV membrane protein (Potential). The
 CC largest part of the protein is cytoplasmic, while its C-terminal
 CC part is associated with the nuclear envelope, most probably the
 CC outer nuclear membrane. Remains associated with the nuclear
 CC envelope during its breakdown in mitotic cells.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=9;
 CC Name=1;
 CC IsoId=Q8WXH0-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q8WXH0-2; Sequence=VSP_007164, VSP_007166;
 CC Note=No experimental confirmation available;
 CC Name=3;
 CC IsoId=Q8WXH0-3; Sequence=VSP_007155;
 CC Note=Produced by exon skipping that results in a frameshift. NO
 CC experimental confirmation available;
 CC Name=4; Synonym=beta;
 CC IsoId=Q8WXH0-4; Sequence=VSP_007156;
 CC Name=5; Synonym=alpha;
 CC IsoId=Q8WXH0-5; Sequence=VSP_007157, VSP_007164, VSP_007165;
 CC Name=6;
 CC IsoId=Q8WXH0-6; Sequence=VSP_007158, VSP_007165, VSP_007166;
 CC Note=No experimental confirmation available;
 CC Name=7; Synonym=gamma;
 CC IsoId=Q8WXH0-7; Sequence=VSP_007154, VSP_007163;
 CC Name=8;
 CC IsoId=Q8WXH0-8; Sequence=VSP_007161, VSP_007162;
 CC Note=No experimental confirmation available;
 CC Name=9; Synonym=NUNCB-N-33;
 CC IsoId=Q8WXH0-9; Sequence=VSP_007159, VSP_007160;
 CC -1- TISSUE SPECIFICITY: Widely expressed, with higher level in kidney,
 CC adult and fetal liver, stomach and placenta. Weakly expressed in
 CC skeletal muscle and brain. Isoform 5 is highly expressed in

```

CC pancreas, skeletal muscle and heart.
CC
CC -1- DOMAIN: The Klarsicht domain mediates the nuclear envelope
CC targeting.
CC -1- SIMILARITY: Belongs to the Nesprin family.
CC -1- SIMILARITY: Contains 1 actin-binding domain.
CC -1- SIMILARITY: Contains 1 calponin-homology (CH) domains.
CC -1- SIMILARITY: Contains 1 Klarsicht domain.
CC -1- SIMILARITY: Contains 1 leucine-rich (LRR) repeats.
CC -1- SIMILARITY: Contains 9 spectrin repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC
CC EMBL: AF435010; AAL33547.1; -
CC EMBL: AF435011; AAL33548.1; -
CC EMBL: AY061757; AAL33800.1; -
CC EMBL: AY061758; AAL33801.1; -
CC EMBL: AY061759; AAL33802.1; -
CC EMBL: AF495911; AAN60443.1; -
CC EMBL: AL117404; CAB55905.1; -
CC EMBL: AL152832; -; NOT_ANNOTATED_CDS.
CC EMBL: AL355094; -; NOT_ANNOTATED_CDS.
CC
Query Match 1.6%; Score 126.5; DB 1; Length 6885;
Best Local Similarity 17.8%; Pred. No. 67;
Matches 219; Conservative 174; Mismatches 408; Indels 431; Gaps 56;
QY 342 PSYNSNYLTRLQKQPKLEVRGAEIRCPDPSPSTVP---TSVRLKPADINIVGA 396
DB 5598 FLYCCEKVIQLLEKIEBALKV-----DVANSIPLEEQKTYKMLEAIVSINQT 5646
QY 397 LGDSLITANGAGSGTFGNLDV-----LTQYRGLS--WS---VGDDENIGTYTLIAN 442
DB 5647 IADSV-----TSLQLDITTEIENRPEFTTFESKILDRQONAVQGRKRGDVGILVR 5700
QY 443 ILREFNPSLKFSVGTGKETSPNAFLNQAAGVGRREDLPVQARRLVDMKNDTRIHFF--- 499
DB 5701 QMOPFTTSVENLFF---RFLTDITSLLSAVVQGEKPSL-YQTRSLIHBLKA-KELHPPRR 5754
QY 500 -----QEDWK-----ITLTFGGNDLCDF 518
DB 5755 RTTCALTLGAGEKLLTDTDLTKTESVGRRI SOLQSDWKMEPQLAEMIKQFSGVETWDQ 5814
QY 519 C-----NDL-----VHSPSPFNFDNIGKALDILHA 543
DB 5815 CEKKIKELKSRLOYLKAQSEDEPLPELHEDLHNEKELIKELQSOIASLWTON-KELQITWKA 5873
QY 544 EVPRAFVVLVTVLEIVNLR-----LYOEKYYVCPMLIRSLCPVLKPEDNSTELATLIE 599
DB 5874 DLTNR-----VLVEDVWLKEQIEHLHRMEDLCLRAVIRK-----QETEDRLNTWV 5921
QY 600 FNKKFQF-----KTHQLIESGRYDTRDEFTVVVQPFEN-VDMPKTS 640
DB 5922 FNEKKKELCAWLQWQENKVLQTDADISIEWMIKYLQKDCMEBINL-----FSENKQLQKMG 5977
QY 641 EGLPDNSFFAPDCFFHSSKSHSRPA-----SALNNMLPEPVQKTTTHKKENKINI 691
DB 5978 DQL-----IKASNKSPRAEIDDKLNKINDRQOHLFDVVGSHVKKLK----- 6018
QY 692 TCPNQVQPEFLRTYKNSMGQHGFWLPCRRBAPSAHPSTVHALRPADIQVVAALDSDSLAG 751
DB 6019 -----ETPFFIQQLDKNSNLSNLTWLA--RISSELS-----KP-----VVYDVCDQDEIQ 6066
QY 752 NGISGKPPDLDPVTVYQYRGLSVSAGGSDLEENVTLIPNILREFNRLTGVAVGSDADNT 811
DB 6061 KRLAEQDQLGRDIE-----QHSAG-----VESVFNICVL-----LHDS 6094
QY 812 NAFINQA-----VPGAKADILMSQVOT--LMQKMDKDHVRVFNEDWK 851

```

[illegible]

or send an email to license@isb-sib.ch).

CC -----

DR D33006; BA11676.1; -

DR EMBL; Z28201; CA82046.1; -

DR EMBL; Z28200; CA82044.1; -

DR PIR; S78475; S78475.

DR SGD; S0001684; MNN4.

DR GO; GO:0006487; P:N-linked glycosylation; IMP.

DR GO; GO:0006493; P:O-linked glycosylation; IMP.

DR GO; GO:0006950; P:response to stress; IDA.

KW Transmembrane; Signal-anchor; Repeat.

FT TRANSMEM 1 28

FT DOMAIN 49 1178

FT DOMAIN 1032 1174

FT DOMAIN 1042 1174

FT REPEAT 1042 1049

FT REPEAT 1050 1057

FT REPEAT 1058 1065

FT REPEAT 1066 1073

FT REPEAT 1074 1081

FT REPEAT 1082 1089

FT REPEAT 1090 1097

FT REPEAT 1098 1105

FT REPEAT 1106 1113

FT REPEAT 1114 1121

FT REPEAT 1122 1129

FT REPEAT 1130 1137

FT REPEAT 1138 1144

FT REPEAT 1145 1152

FT REPEAT 1153 1160

FT REPEAT 1161 1168

FT REPEAT 1169 1174

FT DOMAIN 37 40

FT DOMAIN 1042 1045

FT DOMAIN 1046 1049

FT DOMAIN 1050 1053

FT DOMAIN 1054 1057

FT DOMAIN 1058 1061

FT DOMAIN 1062 1065

FT DOMAIN 1066 1069

FT DOMAIN 1070 1073

FT DOMAIN 1074 1077

FT DOMAIN 1078 1081

FT DOMAIN 1082 1085

FT DOMAIN 1086 1089

FT DOMAIN 1094 1097

FT DOMAIN 1098 1101

FT DOMAIN 1102 1105

FT DOMAIN 1134 1137

FT DOMAIN 1157 1160

FT DOMAIN 1165 1168

SQ SEQUENCE 1178 AA; 139380 MW; BC05DAE0AEFCB282 CRC64;

Query Match 1.68; Score 125; DB 1; Length 1178;

Best Local Similarity 18.08; Pred. No. 4.8;

Matches 198; Conservative 134; Mismatches 317; Indels 448; Gaps 54;

QY 256 SYGAAANSLASRYSEQSFVFPFFYE---TTPSLHSE----- 294

DB 69 SDDDSWYSILTSSFKNDK---IOPAKTLYENLKFGINPKWVNYTTLQNDLLSVKMPRK 125

QY 295 -----DPLQDSTTL-----AMILW----- 309

DB 126 GSKLESYDELKFPDRLTWSVLNLQNNDDADQPEKLPSWSYDWTTPHKLKLSIDK 185

QY 310 -----NMMPEAGKGD---EPLSVKGRPMKCPSESPLYLSYNSNTLTRL---QKPOD 358

DB 186 TVLPICNLFPSAFDKESLEALETIGEP-----LFLYERPKYAQKLMYKRAANOD 235

QY 359 KLEVRREGAEIRRC-----PD-----KQPSDVTVPVSVRLKRPADINVIGAL 397

DB 236 RIKDSKELKKGSKLFTPDGHS PKGLRFNTQFOIKELYDKRBEVYQCAR---NYI--- 290

QY 398 GDSLTGANGAGSPGVNLDVLTQYRLSWSVGDEIGIVTTL-----ANILREF 447

DB 291 -----LTTQSHPLSISIESDSTYQVPLQTEKSKNLVOSGLQER 331

QY 448 -NPSLKFSVGCTKETS PNAFLNQAAGRAEDLPQARBLVDLMKNDRIHQEDMKII 506

DB 332 INDNIN-----STKRRK-----NKQDEVNHRRLQEPVNNQVN-----SLX 370

QY 507 TLEFGNDLCDFCNLDVHYSPOFTDNICKALDIHAEPRAVNLVTYLEIVNLELYQ 566

DB 371 KLEIETDKFTFDKLDVLPSPDFKDAKIE-----ELBS 407

QY 567 EKVYVPRM-----ILRSLCPV-----LKF-----DNDSTELAT 596

DB 408 QKKLYPDKFSAHNENYLNLSKSVKTS PALQRFYFEGAGAVYQKMGHRRDRFPFNVT 467

QY 597 LIEFNKKFOEKTQHLIESGRYDTRDFTVVOFPFENYDMPKTSBGLPNSFFAPDCFH 656

DB 468 LINDKQEQYQARLNSMR-----TFQKT-----KANGI-----I 496

QY 657 SSRSHRAASALMNNLEPVGQTTTRHKEENKINITCPNOVQPLFTYNS-----NQ 709

DB 497 SWLSHGTLGYLYNGMAFP-----WDNDFDLQMPKHLQLLSQFYQSLILEDPQ 547

QY 710 GHGTWLPCKDRAPSAHPYSVHALRPADIQVAAALGDSILTA---GNGISKPPDLDVYT 766

DB 548 GNGRYF-----LDVSDSLTVRINGNGKNNDARFIDVPT 581

QY 767 QYGLGYSAAGDGLNVTTLPNILREFNRNLGYVGTGADNDTAFNQAAPGAKAD 826

DB 582 ---GLYI-----DITGLASTAPSRD---YANSYEEELQBE 612

QY 827 LM-----SQVQTLQMKMDHRVNFHEMDKVTTLIGSDLCDYCT-DSN-----L 871

DB 613 HLDINNIPESNGETATLPDKV-DDGLVM-----ATLITELADYITSDENKHKRY 663

QY 872 YSAANVDHRLNLDVLRREVPRVLYNVDLPNTIMRQVFLNPKCPVQAQASVLCNCV 931

DB 664 PTDIDKDLKLELE---ELPK-SKTIENKLP---KQRYFLEKXK-----LYNC- 707

QY 932 LTLRENSQELARLEARSRAVRSRMEIAGSGRYDQEDSFVVLPFFONITQLEVLADGP 991

DB 708 ---RNN-----HNSFEELSPINTYFHGV---PAL---IP 734

QY 992 DTSPFAPDCIHPNQKHSQALRALWTNMLEPLGSKTETLIDRAEMPITCPTONEPFLRTP 1051

DB 735 HRRTY---CLHNEYHVPDRYAFDAYKN-----TAYLDEPFWFYDGLKIC 777

QY 1052 RNSNYTPYIKPAIENMGSDPL---CTEWKA---SNSVPT-SVHOLRAPADIYV----- 1097

DB 778 SNINSWYPPNIPINSWNPMLKLEISSTKPEKSLFDSNKVSEYFKNLSMDVDVLIKNIP 837

QY 1098 -AALGDSLTAVAGARPNNSSDLPSTSRGLSWSIGCGDNTL-----THTTLPNILKFF- 1148

DB 838 KAFIEVFTLYL-----NSENV-TAYRQKELEIYQCNLTFLIEKKLLHQLRINVA PKLS 891

QY 1149 ---NPLYLGFSTTWE 1161

DB 892 SPAKDPFLGGEKAMWK 908

RESULT 10

YFAS ECOS7

ID YFAS ECOS7 STANDARD; PRT; 1534 AA.

AC OSKES5:

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical protein yfas precursor.

GN YFAS OR Z3481 OR ECS3111.

OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia..
 ON NCBI_TaxID=83334;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / BDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
 RL Nature 409:529-533(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952;
 RX MEDLINE=21156211; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shingawa H.;
 RA "Complete genome sequence of enterohaemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 CC -1- SIMILARITY: BELONGS TO THE UPF0192 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AE005454; AAG57357.1; -
 DR EMBL; AP002560; BAB36534.1; -
 DR PIR; G91017; G91017.
 KW Hypothetical protein; Signal; Complete proteome.
 FT SIGNAL 1 38
 FT CHAIN 39 1534 HYPOTHETICAL PROTEIN YFAS.
 SQ SEQUENCE 1534 AA; 169893 MW; E326FD052770F1A3 CRC64;
 Query Match 1.6%; Score 123.5; DB 1; Length 1534;
 Best Local Similarity 18.5%; Pred. No. 9.3;
 Matches 274; Conservative 179; Mismatches 554; Indels 475; Gaps 66;
 QY 161 QPDMKLIINFPSNASCYCLCPASQNGLA-----AGVDLMGVLDLQOE 206
 DB 9 QPDMH-L-SRFGSALACLSLSTVGTLNADSLPSSNYPAPGATPFLDASSFSSSB 67
 QY 207 VPRAFVNLVDLSEVAESRQYH-----GTWLSAPPEPCNCEETRLAKLVMMQMSY-- 257
 DB 68 ESK-----VRLAPRGDYRIRYQMEYGVYVRLYRIPDPAFLRQCKLHRIYVQPOLG 122
 QY 258 -----QEAWSLASSRSESEFTVVPQPFYETTSLSHSEDRLODSTTLAMHLMN 310
 DB 123 DGLNLTLMYLMWDWYKSRVRVQRTFSQSRQNVTOALPEL-----QLGN 167
 QY 311 RMMEPRPG-EKDELSVYKGRPMKCPQSESYLFSYNSNLTQLQKPODKLEVRGAEIR 369
 DB 168 AAIKPRYVONNQS-----PLKRYPLVQGFYPLMQAKPVEPQGVKL-EGA--- 214
 QY 370 CEDKQSDVTPVTSVHLKPADINVIGALGSLTAGAGAGSTPGNVLDVLTQYGLSMSVG 429
 DB 215 -----SSNFISPGNIIYIPLGQ----- 232
 QY 430 GDENICTVTTLANILREPNLSLKGFSVGKETSPPAFLQAVAGRAEDLPVQARLYVD 489
 DB 233 -----QEPGLYIVEAMVGG-----YRATTIV 252

QY 490 LMKNDTRINFQBDWKIITLFIGNDLCD-----FCNDLVHYSQNFQNDIGKALDIH 542
 DB 253 VVSDTVLALSKVSGKELVMTAGKKGGAKPSEILMTDGLGVMTRGVTDSDG-TLQLGH 311
 QY 543 AEVPRAFV-----NLVTLLEIVNLRELQYQEKVYPPRITLSLCCVLFKFDNSTELAT 596
 DB 312 ISPERSYIILKDAEGGVFVSE-----NPFYSEBIYVTRLYIFTRDLRYAGGRVDVKV-- 364
 QY 597 LIEFNKKFOEKTHQ-----LIESGRYDTRBFTVYVQPFENVMPKTSSEG---LPDSFFA 650
 DB 365 ---MGRHFDPHLSSEIYAPAPALSVLDANGSLQTVDVTLTDARRNGQGSFRLPEPAVAG 421
 QY 651 PDCFHFSKSHSRASALNNMLER---VGOKTRHKFE-----NKINITCPN----- 695
 DB 422 GYELRLAYNNQYSSSFRVANYIKHPELGLALDKKEFTGAEVSKQLQLYVDGSPVKN 481
 QY 696 -QVQFLRTYKNMSQCHGTWLPQRDRAPALHPTSVHALRPADIQVVAALGDSLTAGNGI 754
 DB 482 ARVQLSLRAQQLSMVGN-----DLRYAGRPFVS-----LEGSEIVSDAS 520
 QY 755 GSKPDOLPVTTOYQGLSYSAGDGLSENVTLPLIREFNLTNGYVGT---GDAND 810
 DB 521 GHVALNLPPADKPSRYLTVSASDGAAYRVTTKEILIE-RGLAHYSUSTAAYSNSGE 578
 QY 811 TYAFLNQAVPGAK-----AEDLMSQVOTLMQMKDHRVNF-----HED 849
 DB 579 SVVFRYALESKQVRYVTEWRLERDTRSHSGDLPSGGR-SFTVYNDKGNVNLTLRDKD 637
 QY 850 WKVITVL-----IGSDLCVCTDSNLYSANFVDHLRNALDVLREVRVLNVLV 900
 DB 638 GLILAGLSHAVGKSGMSHTGVVDIADKTLVQPG-----ETAKMLITFP 682
 QY 901 DPLNPTIMQVQLGNPDKCPVQASVLCNCV--LTLR--ENGQELARLAFSAVYSSNR 956
 DB 683 EPIDELAL-----TLERDVEQQLSLSPANWLLQLRLNDQYKARVY-- 726
 QY 957 ELVSGRYDTPQDFSVVL-----QPFQNIQLPVLADGLPDTSFAPDCIHPNOKFHSOLA 1012
 DB 727 ---SNSFAPNITTSVLYTRNQYISQNGIKVAVQL-----DIRKDKTHYQPG 774
 QY 1013 RALMTNMLEPLIGSKTETTLDR-----AEMPTICPTQNEBFLTPRNSNYT----- 1057
 DB 775 ELVNVETLSLKGKPPSAQLTVGVDEMIYALQPEIAPNIGKFFYPLGRNNVRTSSLSLF 834
 QY 1058 -----YPIKPAIEMWGSDFLCTEMKASVPSYHQ-LRPADIKY-VAALGSLTT 1106
 DB 835 ISYDQALSEEPVAPG-----ATNRSERRVWLEPRREEVDYTAAMPSSLTT 880
 QY 1107 AVGARPNNSDLP--TSWRGLSMSIGDG-----NLEHTTLPLILKKFNP 1150
 DB 881 DKQKAYFTFLMPDLSLTRITARGNNGDLVGQGRAYLRSEKNLYMKMSMPTVY----- 935
 QY 1151 YLLGFSSTWEGTAGLVNABGARADMPAQMADLVERMKNSPDINLEKDWKLVTLFIGV 1210
 DB 936 -----RVGDKPSAGLFIISQQDNB-----VALVTFAG- 964
 QY 1211 NDLCHYCENPEAHATEVY---OHIOQALDIISEELR-----APV-ANVEVM 1254
 DB 965 ---AEWRQTLTLHGKANYISLAQNIQGS-GLISABEIQQNGVQODSISTKLSEFVDSMPVE 1020
 QY 1255 ELASLYQGGGKGCAMLAQNNCTCLRHSSSLEKSEKELKVN-----W---NLQHG 1302
 DB 1021 QOKNVMLGGGDNALMLPEQASIRL---OSSETPOEIPFNNDALVDPEMGVYINGSR 1077
 QY 1303 SSFSY-W-----HOYTQREDFAVVQPFQNTLTPLNBERG-DTDLTFPSEDCGFHPSDRGH 1355
 DB 1078 IFLSLAMRSLADHQSAANDIRQMIQ--DNRLRLMLQLGAPGARFTWMGED-----GN 1127
 QY 1356 AEMALALNNMLEPVGRKTSNNFTHSRAKLKCPSESESYLTLLNSRLLPQDAEAP-- 1413
 DB 1128 GDAFLTAW-----AMVDAQSQAQVGTQPE--YMOHMLDS--YAOQADNMPPL 1173
 QY 1414 ---EVLVMA---VPVAGVGLVVGIIIGTVVRCRGRGRED 1448

Db 1174 HRAVLAMAOENMLPCKT---LTKGLDEAL---ARRGKTED 1209

RESULT 11

GCN1_YEAST STANDARD; PRT; 2672 AA.

AC P33892;

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Translational activator GCN1.

GN GCN1 OR YGL195W OR G1318.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.

OX NCBI_TaxId=4932;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=93268304; PubMed=8497269;

RA Marton M.J., Crouch D., Hinnebusch A.G.;

RT "GCN1, a translational activator of GCN4 in Saccharomyces cerevisiae, is required for phosphorylation of eukaryotic translation initiation factor 2 by protein kinase GCN2."

RL Mol. Cell. Biol. 13:3541-3556(1993).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=5288C / FY1679;

RX MEDLINE=87197971; PubMed=9046087;

RA Coglievina M., Klima R., Bertani I., Delneri D., Zaccaria P., Brusch C.V.;

RT "Sequencing of a 40.5 kb fragment located on the left arm of chromosome VII from Saccharomyces cerevisiae."

RL Yeast 13:55-64(1997)

RT FUNCTION: TRANSLATIONAL ACTIVATOR OF GCN4. MAY BE INVOLVED IN SENSING CHARGED TRNA AND STIMULATING THE KINASE ACTIVITY OF GCN2 IN AMINO ACID-STARVED CELLS. REQUIRED IN VIVO FOR THE PHOSPHORYLATION OF EIF-2-ALPHA ON SERINE-52 BY THE PROTEIN KINASE GCN2.

CC -1- SUBUNIT: COMPONENT OF A HETEROMERIC COMPLEX THAT INCLUDES GCN1 AND GCN20.

CC -1- SIMILARITY: STRONG, NO S. POME SPACT1866.05C.

CC -1- SIMILARITY: Contains 20 HEAT repeats.

CC -1- SIMILARITY: Contains 20 HEAT repeats.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.ebi.ac.uk/announcements> or send an email to license@sb-sib.ch).

CC -----

EMBL: L12467; AAA34635.1; -

EMBL: X91837; CAA62949.1; -

EMBL: Z72717; CAA96907.1; -

PIR: A48126; A48126.

SGD: S0003163; GCN1.

DR GO: G0005830; C:cytosolic ribosome (sensu Eukarya); IDA.

DR GO: G0006448; P:regulation of translational elongation; IMP.

DR InterPro: IPR000357; HEAT repeat.

DR PROSITE: PS50077; HEAT REPEAT; 4.

DR Translation regulation: Activator; Repeat.

FT REPEAT 932 970 HEAT 1.

FT REPEAT 1030 1067 HEAT 2.

FT REPEAT 1099 1138 HEAT 3.

FT REPEAT 1243 1281 HEAT 4.

FT REPEAT 1284 1321 HEAT 5.

FT REPEAT 1405 1442 HEAT 6.

FT REPEAT 1484 1521 HEAT 7.

FT REPEAT 1523 1559 HEAT 8.

FT REPEAT 1561 1598 HEAT 9.

FT REPEAT 1603 1640 HEAT 10.

FT REPEAT 1641 1679 HEAT 11.

FT REPEAT 1721 1758 HEAT 12.

FT REPEAT 1760 1796 HEAT 13.

FT REPEAT 1862 1903 HEAT 14.

FT REPEAT 1905 1942 HEAT 15.

FT REPEAT 1947 1984 HEAT 16.

FT REPEAT 1985 2024 HEAT 17.

FT REPEAT 2097 2134 HEAT 18.

FT REPEAT 2290 2328 HEAT 19.

FT REPEAT 2347 2384 HEAT 20.

SEQUENCE 2672 AA; 296693 MW; 980FPD0375359DLC CRC64;

Query Match 1.6%; Score 123.5; DB 1; Length 2672;

Best Local Similarity 19.9%; Pred. No. 23;

Matches 213; Conservative 139; Mismatches 339; Indels 381; Gaps 53;

Qy 165 KLINVFSSN-----ASQCYLCPASQONGLAAGVDELMGVLLYDQEEVP 208

Db 1447 KLIPVAVSNIDEIAMRTKGSVQLLGNMAYLDPT-QLSASLSTIVPEIVGLVDSHKEVR 1505

Qy 209 RAFVNLVDSSEVAEVRQYHGTWLSAPAPCNCEETTRIAKYVMQMSYQEAANSLASS 268

Db 1506 KAADF-SLKRFEVIR-----NPEIQKLVPLVLC-----ALGDPT 1539

Qy 269 RYSEQESFTVVFQPF-FYETTPSLHSEDPRLDSTTLAMHLNRMMEPAGEKDEPLSVKH 327

Db 1540 KYTEBALDSLIIQTFVHYIDGPEL-----ALIIHIIHR-----GMHDSANIKR 1583

Qy 328 GRMKCPSEQSPFLFYSRNSNYLTRLOKQDKLEVE-----GAEI-RCP 371

Db 1584 ---KCKIKVGNNAIIVDTKDLIPYLQIDEVEIMVDEPENTRATARALALVERLG 1639

Qy 372 DKDPSTVPTSVARLKPADINVGALDSLTAGAGSTFGNVLDTQYRGLSWSVG- 430

Db 1640 EEGFPLIP-----RLIDTLDSEKSGDRGSA-----QALAEVLSGL 1677

Qy 431 ---DENIGT---VTLANIIRE-FNPSLKFSYGTGKESIPNAPLN---QAVAGRA 477

Db 1678 GLTKLEMLPTLIAGTNGFRAYIREGFMPFLPLPCVCGQFAP-YINQIIPILSGLA 1735

Qy 478 E-----DLPVQARL-----VDLM-----KND----- 494

Db 1736 DNDENIRDTALKGKILVKNVATKAVDLPILPERGMFDENDIRLSSVQLTGEELLFOVT 1795

Qy 495 ---TRIHFOEDMKITLFIGNDLCPNDLVHVSFQNFDTNIGKAL----- 538

Db 1796 GISSRNEFSE-----DGDNGEFSGLVVDLQDRDRRLALFLVGRNDTSIVRA 1847

Qy 539 ---DILAEV---PRAFNLVTLVLEIVNLRELVOEKVYCPRMILRSLCPVLKPDNS- 591

Db 1848 TVVDIWKALVPNPRAVKEILPTLTGMIVTHLASSNV-LRNIAQTLGLVRRVGNAL 1906

Qy 592 TELATLIEFKKQKQKTHOLIESGRYDREDFVVVQPFPEVNDMKTS- 641

Db 1907 SOLPLSLE-----ESLIFRNSDSRGVCIALYELIESASTETTSQFQSTIVNIIR 1957

Qy 642 -GLPDNSFPAPDCFHRSKSHRASALMNMMLPEVGQKTRRKFENKINITEPNOVQPF 700

Db 1958 TALIDES-----ATVEEAALSFDV-----FQDVVGKTAIVDELTPY 1993

Qy 701 LRTYKSMQSGHTWLPDRAPAPSLHPTSVHALRPADIOVAALGSLTAGNGISGKPD 760

Db 1994 L-----LHMLESSDNSDFALLG-----LQGISMSKSD 2020

Qy 761 L-----PDVTQYRGLSYSGAGSGSLENV---TTLPNLIREFNRLTGVAVGTGD 807

Db 2021 VIPPIIIPITLAPPID-----AFRASALGSALAEVAGSALYKSLSTIINALVDALITGS 2074

Qy 808 ANDTNAPFLNDAVPGA-----KAEDLMSOVQTLNQKKDDH---RV-----NFHEMDKV 852

Db 2075 DESTKALBELALRVFLSVNVDGLHPLAQIISLILKSDNIKRIAVLERLPFPDK--- 2131

Qy 853 ITVLIGSGDLCDYCTDSNLVSAANFVDHLRNALDVLHREVPRVLVNLVDFLNPITMKQVF 912

Db 2132 -TUL-----DEVT-IPNFVSHALLSLD---DEDQVNVNGFNALS-TLKKVD 2174
 Qy 913 LGNDK-PCVQASVLC-----NCVLT-----RENSOELA 942
 Db 2175 KPTLEKVKRAKOSLALTGOCODVAFKLPRGNCVLPFLGIMVGNDEBERE-ALA 2233
 Qy 943 RLEAFBRVAVRSNRELVGSGRYDTEQDFSVLQPFQCNQLPYLADGLPDTSEFA----- 997
 Db 2234 IADVSKTPAANKPFVS---VITGPIRVGGERSSDIKAIL-----FALNVLF 2281
 Qy 998 ---PDCIHPQKHSQALRYALNTMELPSKTEITLDLRAEMPTCTQNEP 1046
 Db 2282 IKIPMLRP---FIPOLQRTFVXSLD---ATNETTLRLAAKALGALIEHQ 2327
 RESULT 12
 PDP2_ARATH STANDARD; PRT; 1039 AA.
 ID PDP2_ARATH STANDARD; PRT; 1039 AA.
 AC 09M9W8;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Phospholipase D p2 (EC 3.1.4.4) (AcPLD2) (Phospholipase D2 PHOX and
 EX containing domain) (Phospholipase D zeta 2) (PLDzeta2).
 GN PLD2 OR AT3G05630 OR F18C1.10.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
 NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016720; PubMed=11130713;
 RA Salanoubat M., Lemcke K., Rieger M., Ansoerge W., Unseld M.,
 RA Fartmann B., Vailly G., Bloeker R., Perez-Alonso M., Obermayer B.,
 RA Delgany M., Boutry M., Griwall L.A., Mache R., Pilsdomech P.,
 RA De Simone V., Cholase N., Artiguenave F., Robert C., Broetier P.,
 RA Wincker P., Catolico L., Weissenbach J., Saurin W., Querrier F.,
 RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
 RA Wurmach E., Dizonek H., Erlie H., Jordan N., Bangert S.,
 RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
 RA Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simionati B.,
 RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordstiek G.,
 RA Reichelt J., Scharte M., Schoen O., Barynes M., Terol J., Clement J.,
 RA Navarro P., Collado C., Perez-Perez A., Ottensmeyer B., Duchemin D.,
 RA de Haan M., Maarte A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
 RA Monfort A., Argilou A., Flores M., Liguori R., Vitale D.,
 RA Manhaupt G., Haase D., Schoof H., Ruid S., Zaccaria P., Mewes H.-W.,
 RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
 RA Rooney T., Rizzo M., Walters A., Utecherback T., Fujii C.Y., Shea T.P.,
 RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
 RA Pal G., Miltcher J., Sellers P., Gill J.E., Feldblum T.V.,
 RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
 RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asanizu E.,
 RA Sasano S., Kimura T., Ideasa K., Kawashima K., Kishida Y.,
 RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
 RA Watanabe A., Yamada M., Yasuda M., Tabata S.,
 RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
 thaliana."
 RT Nature 408:820-822(2000).
 RL Nature 408:820-822(2000).
 CC -1- FUNCTION: Hydrolyzes glycerol-phospholipids at the terminal
 CC phosphodiesteric bond. Phosphatidylcholine-selective.
 CC -1- CATALYTIC ACTIVITY: A phosphatidylcholine + H(2)O = choline + a
 CC phosphatidate.
 CC -1- ENZYME REGULATION: Calcium-independent and PIP2-dependent.
 CC -1- SIMILARITY: Belongs to the phospholipase D family. PXP-PLD
 CC subfamily.
 CC -1- SIMILARITY: Contains 1 PHOX homology (PX) domain.
 CC -1- SIMILARITY: Contains 1 PH domain.
 CC -1- SIMILARITY: Contains 2 PLD phosphodiesterase domains.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL: AC011620; AAP26134.1; -
 DR InterPro: IPR001849; PH.
 DR InterPro: IPR001736; PLD.
 DR InterPro: IPR001683; PX.
 DR Pfam: PF00169; PH; 1.
 DR Pfam: PF00614; PLDC; 2.
 DR SMART: SM00233; PH; 1.
 DR SMART: SM00155; PLDC; 2.
 DR PROSITE: PS50003; PH DOMAIN; 1.
 DR PROSITE: PS50035; PLD; 2.
 DR PROSITE: PS50195; PX; FALSE NEG.
 KM Hydrolyase; lipid degradation, Repeat; Multigene family.
 FT DOMAIN 45 205
 FT DOMAIN 215 343
 FT DOMAIN 465 492
 FT DOMAIN 840 867
 FT ACT_SITE 470 470
 FT ACT_SITE 472 472
 FT ACT_SITE 477 477
 FT ACT_SITE 845 845
 FT ACT_SITE 847 847
 FT ACT_SITE 852 852
 SQ SEQUENCE 1039 AA; 117902 MW; A8433C237C3B7F73 CRC64;
 Query Match 1.6%; Score 123; DB 1; Length 1039;
 Best Local Similarity 18.7%; Pred. No. 5.4;
 Matches 157; Conservative 123; Mismatches 313; Indels 246; Gaps 37;
 Qy 412 GNVLDTLT-QYNGLSVSGDENIGVTTTLNLRBFNSLKGFSVGTGKET----- 462
 Db 273 GKLDLWPTDTGLQ---GTSESSQPRRLAEQVKEHNPRLRFQKVTSGDRTVRLRTSS 328
 Qy 463 -----SPNAFLNQAAGRAEDLPVQARLYD-----LMKNDTR 496
 Db 329 RYKKEVAVKADGACGSHRFSFAPRGLTSD-GSQAMFWDGHAFALMAGNATS 387
 Qy 497 IHFQEDWKITTLFIGNDLDCFDNDLVHVSPOFTNDNICKALDILHAEPRAFNVLTVL 556
 Db 388 EIFMTGWWL-----CPRL-YIKRPFEDHPSLRDLALLETAKQGVKEVQI- 431
 Qy 557 EIVNLELVOEKVYCPRMILSLSCVYKFPDNLSTELATLIEFNKKPQEKTHQL----- 611
 Db 432 -AKINSLYSKRLQNIHKNV---VLRYPHLSGIGLYLSNHEKIVYVDYQVCFIG 485
 Qy 612 --IESGRYTRDPTVVQVPPF---ENVMPKTSBGLPNSFPAPCPHFSSKSHR--- 663
 Db 486 LDCFRKYVTAHKIGDCPRYIWPCKDYNNPESB---FNSWETMKDELDRKKYRMPV 542
 Qy 664 --AASALNNMLBPVQOKTRHKFNKINITCPNQVQPLRTYKNSMOGHVLPQRDRA 721
 Db 543 HDVHCALMPRCRDVARRHVFQWNSHKRKAPEQETIPLMHHNVLP- YLGTRE-- 598
 Qy 722 PSHLHTSVHALRP--ADIVVAALGDSLTAONGSGRPDLDPVTTQYRGLSYAGCD 778
 Db 599 ----IDITIAAKPPEEDPKPVLLAHDSFSS---ASPPQELPILLPETDAPFAGRD 649
 Qy 779 -----GSLKVVTLPLILREF-----NRNLGVAVGTDAADT-- 811
 Db 650 LKUDSGARODPGETSESULDEAVNDMMWQIGKQSCRCQIIRSVQWAGSISQPDSTH 709
 Qy 812 -----NAFLNQAVGARAED--LMSQVOTLMQKDKDHRVNFHEWKV 852
 Db 710 RAVCSLIGNAHEFIYENQF---ISGLEKEDTILNRVLEALYRRI-----LKAHEENKC 761

QY 853 ITVL1-----GGSDLCDYCTDNLVSANFVDHRLNALDYLHREVPVLNVLDFL 903
 Db 762 FRVIVIPBLPGGGID-----DFGAT-----VRALMHWQYTIIRREGISINDL 808
 QY 904 N-----PTMRQVFLN-----PDKCPVOQASVLCNCVL 932
 Db 809 NALLGPKTDYISFYGLRSYGRLFEDGPATISQIVHASKMIVDRIVAGSSINIRSL 868
 QY 933 TLRENSOELARLE--AFSRAVRSRMRVLGSGRYDTQ-----EDPSVVLQPFQFQNIOLPV- 985
 Db 869 LGRSDSEIGVIVIEDEKFEVESNMGKMMWAGKFSYLSGCSLMSHILGLHAGEIQKTEPFIK 928
 QY 986 -----LADGLPDTSPFAP--DCIHPNOKFHSQALRALMTWML-----EPYGSKTETLIDL 1032
 Db 929 DATYKOLMMAATAKNTDINQVFSCL-PNE--HIRSRPALRHNNALCKDKLGHITTDLGI 985
 QY 1033 RAEMPTCPQNEBFLTPPNNSNTYVPIKPAIENWGSDFLCTEWMKSNVPTSVHQLRP 1091
 Db 986 APERLESQGSWSWELTKETRGNIWCPFLQ-----FMCDQ-----EDLRP 1024

RESULT 13
 GALK_SCHPO STANDARD; PRT; 713 AA.
 AC Q9HDJ3;
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE GAL10 bifunctional protein [includes: UDP-glucose 4-epimerase
 (EC 5.1.3.2) (Galactowaldenase); Aldose 1-epimerase (EC 5.1.3.3)
 (Mutarotase)].
 GN GAL10 OR SPBPB2B2.12C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetaceae; Schizosaccharomycetaceae;
 OC NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N. A.
 RC STRAIN=972;
 RX MEDLINE=21848401; Pubmed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Scouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brookes K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hildalgo J., Hodgson G.,
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moulis S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Shelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voickert G., Aert R., Roben J., Grymopiez B.,
 RA Welfens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabell C., Fuchs M., Fritz C., Holzer E., Moesli D., Hilbert H.,
 RA Beyer P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu S., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore C., Hurst S.M.,
 RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Rhode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey P., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Uesery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC -!- FUNCTION: Mutarotase converts alpha-aldose to the beta-anomer. It
 CC is active on D-glucose, L-arabinose, D-xylose, D-galactose.
 CC -!- CATALYTIC ACTIVITY: UDP-glucose = UDP-galactose.
 CC -!- CATALYTIC ACTIVITY: UDP-glucose = UDP-galactose.
 CC -!- CATALYTIC ACTIVITY: Alpha-D-glucose = beta-D-glucose.
 CC -!- COFACTOR: NAD.

CC -!- PATHWAY: Galactose metabolism; third step.
 CC -!- PATHWAY: Hexose metabolism.
 CC -!- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE SUGAR
 CC EPIMERASE FAMILY.
 CC -!- SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO THE ALDOSE
 CC EPIMERASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL: AL512522; CAC21414.1; -.
 CC HSSP: P09147; 1XEL.
 CC GeneDB: Spombe; SPBPB2B2.12C; -.
 CC InterPro: IPR001823; Ald1 epimerase.
 CC InterPro: IPR001509; Epimerase_Dh.
 CC InterPro: IPR005886; GalE.
 CC InterPro: IPR000205; NAD_binding.
 CC Pfam: PF01263; Aldose_epim; 1.
 CC Pfam: PF01370; Epimerase; 1.
 CC TIGRFAMs: TIGR01179; galE; 1.
 CC PROSITE: PS00545; ALDOSE_1_EPIMERASE; 1.
 CC KEGG: Multi-functional enzyme; Isomerase; NAD; Galactose metabolism.
 CC FT DOMAIN 1 350
 CC FT DOMAIN 351 713
 CC FT NP_BIND 7 38
 CC FT ACT_SITE 532 532
 CC SEQUENCE 713 AA; 80666 MW; 51C89DA0843A8556 CRC64;
 SQ

Query Match 1.6%; Score 122.5; DB 1; Length 713;
 Best Local Similarity 19.8%; Pred. No. 3.2; Indels 231; Gaps 38;
 Matches 138; Conservative 92; Mismatches 235;
 QY 370 CPDLPSPD-----TVPTSVHRLKPAD-----INVIGALDSTLAGAGSTPG 412
 Db 149 CPN-DPTNPGKTKVAIENIKDLHTSDMTWRGAILRYNPICAHSGGLGEPLG-IPN 206
 QY 413 NVLDVLTQY--RGLSMVQGD--NIGT-----VTTLA-----NIREFNP-- 450
 Db 207 NLPLPLAQLVAGREKLVFGDDYDSHDPRIIDYIHVVLDLAKGHIALNLYMKINNSEG 266
 QY 451 -LKGFSVGTCKETSP-----NAPLNAVAGRADLP--VQARR--LVLDKMDTTHIQ 500
 Db 267 MYREMNIGTCKGSSVFDIYHAFCKEV-----GDLPEYVAGRTGVDVNLITASPNRANSE 321
 QY 501 EDKITTLLFGNDLDCDFCNDLVHSPNFTDNIGKALDILHAEVPRAVNLVTVEIYN 560
 Db 322 LKKK-----AEISTDACHDLMKWTEN-----PFGF-----QIDN 352
 QY 561 LR-ELYOEKKVYCPMILRSICPVLPKPDNSTELATLIEFNKFOEKTHQLI-----ESG 615
 Db 353 YKKKLTFTLIGMVKRLHTIC-----FQDLVSIANYGALVQAVYKGRNLVNGFNDFS 407
 QY 616 RYTRFEDFTVVOVPPFENVDMPKTSBGLPDNSFFADPCFFSKSKSRASALMNNMLP 675
 Db 408 RYKLKEN-----PFF-----GATIGFRAMRIANQOFEV 435
 QY 676 VQKTRHKPE-NKINITCPNQ-----VQPELRTYK--NSMQ-----GHGTWLPGR 718
 Db 436 DGHLYTLCKENKNTTLHGNGNCFDQFGLGARQYEDYNTLFLVLDKGNNGF-----491
 QY 719 DRAPSAALPHTSVVALRPADIQVVALGDSITAGNGIGSKRPDLVDVTTQYRGLSYAGGD 778
 Db 492 ---PSDLFTLVKTKITNNLSLEI-----EKSVIPEYSKL-----522
 QY 779 GSIENVTTLNLIREFNRNLTGYAV-----GTGDANDTNAFLN-QAVPGAKAED 826
 Db 523 ---NTAV-----NLTNHSYNNLASPNKTIIDGTIIKSTYNNYLKVNSSETSPGTG 569

QY 827 LMSQVOTLMQKMDHRVNEHEDMKVITVLIGSSDLCVCTDSNLYSANFVHLRNALD 886
 DB 570 IYEQMDITKPTLDIPISFD-----NCRTVDREASKFCLDTFRKS-----LKNIVE 616
 QY 887 VLRREPRVLYNVLDLNPTRIMQVFLGN-PDKCPVQASVLCNCLTLRENSQELARLE 945
 DB 617 VLRPSVPLKLV-VSTTEPAF--QLYTGDDNDICEQSRSGFC-----VETGRFT 662
 QY 946 AFSRAVSSRRELIVSGSR-YDQEDFSVVLQPFPPON 980
 DB 663 NALNNEKSKQVILRKGEVYGARSKFSLVADLEEN 698
 RESULT 14
 SC71 SCHPO STANDARD; PRT; 1811 AA.
 ID SC71 SCHPO STANDARD; PRT; 1811 AA.
 AC 09UT02; 014168; 42, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Protein transport protein sec71.
 GN SEC71 OR SEC7A OR SPAC4D7.01C OR SPAP8A3.15C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetaceae; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=2148401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Baaham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Wolckstein G., Aert R., Robben J., Grymoprez B.,
 RA Weljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Holzer H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallardin C., Talada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revelante J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RA "The genome sequence of Schizosaccharomycetes pombe."
 RT Nature 415:871-880(2002).
 RL Nature 415:871-880(2002).
 CC -1- FUNCTION: May play a role in protein transport (By similarity).
 CC -1- SIMILARITY: Contains 1 SEC7 domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.ebi.ac.uk/announce/
 CC or send an email to license@ebi.ac.uk).
 CC EMBL: AL117210; CAB55182.1; -
 DR EMBL: 298602; CAB11286.1; -
 DR PIR: T39252; T39252.
 DR HSSP: Q99418; 1PBV.
 DR GeneDB: SPombe; SPAC4D7.01c; -

DR InterPro: IPR000904; Sec7.
 DR Pfam: PF01369; Sec7; 1.
 DR SMART: SM00222; Sec7; 1.
 DR PROSITE: PS0190; Sec7; 1.
 DR TRANSPORT: Protein transport.
 FT DOMAIN 692 880
 SQ SEQUENCE 1811 AA; 206930 MW; 14AC990B746804D CRC64;
 Query Match 1.6%; Score 122; DB 1; Length 1811;
 Best Local Similarity 17.4%; Pred. No. 15;
 Matches 284; Conservative 234; Mismatches 523; Indels 594; Gaps 79;
 QY 19 GTPQHTSPKSTL-EGQLMPETLKN-SPPCPNPKLVNMSKSVHSLKPSDIKVAAT 76
 DB 325 GTPDRANSPIPOLSKNKLTLSEFHRKSPDQVREAPLEBDSLEQQLRDAFL-LIRAL 383
 QY 77 GNLRIPPDGTGDLKQDTERPQVCMV-----MTVLSDI-IRYFSP-SVMP-- 124
 DB 384 CKLSIKNIPIYEHYDLKQSMRSKLSLHLYHILRTVMNLIISLDIVNKIRSPSTPTPLI 443
 QY 125 -----VC-HTKRVPFH-----DGAEDLMTQAQELVNMKNENQLDF----- 160
 DB 444 DAVKQYICLALAKNVSHVLPVFEISCEIFWLILSLKNFPSESEVFTETLFPPILEMR 503
 QY 161 -----QFDMKLIVNFFSNASQ-----CYLCPASQONGLAAGVDELMGVLDVLAQEV 207
 DB 504 TSSNQOKYVLNLIHFRMCSEPQTLIELYNYDICSNTENIYERAILVTLISRIASQSTSP 563
 QY 208 PRAFVNLYD-----LSEVAEVSROYHGTWLSPAEP-----CNCS 242
 DB 564 PPSFVPRDQVLIDKPGFYHTLNDIPQLNASTIGSYVSHNPPYPDYQIRLKSVCRLIS 623
 QY 243 EETRLAKVNMQSVQANSLASGRSEQSSFTVPDPFFETTPSLHSEPRQDST 302
 DB 624 TLSS-----LFTWCQTFAPVETAKDQDTESTSGEPP-----QKXSEPPSAGINS 672
 QY 303 TLAMHLNMRMPAGE---KDEPLSVKHGRPMKCPQSEPFYFSYNSNVLTRLOKPOK 359
 DB 673 TSMNVL-----SSGQALATDDPSQEPENKHKKQQLQEAIOKFN-----KP--- 714
 QY 360 LEVRGAEIRCPDKDPSDTVPSTVHRLKPADINVIQALDLSLJAGNAGSTPGNVLDVLT 419
 DB 715 -----KEGKILLSHFFASTKPTDIKPF-----LSTEGDLKAVIG 751
 QY 420 QYRGLMSVSGENIGTVT-----LANIRENPSLKGSVGTGKETSNAVL 468
 DB 752 EYLG-----EGNDENAIHMSFVDHMSFNDIPRVNMLRSP---LQKR----- 791
 QY 469 NQAVAGGRAEDLPVOARRLVLDLMDKQDTRHNPQEDWKIITLFIGNDLQDFCNDLVHYSPO 528
 DB 792 -----LPGEKQIDRFM-----LKFARK 809
 QY 529 NPTDNIGKALDILHAEVPPAFVNLVTVLEIVNLRELYOKKVCPRMILRSLCPVLKPD 588
 DB 810 YIDDNLG-----VFNQADATVILAVSIIML-----NTDLHSPQVKRMTCQDPIKON 856
 QY 589 ---DNSTELAT-LIIPNKKPQEKTHQLESRGYRTREDPFTVAVQPFENVDMKTSSEGL 643
 DB 857 RGVDGAGNISDSFLEVEVEIOK--NEIVLKDQEOPTSNFP-----EIPGTS--- 901
 QY 644 PINSFPAP-----DCHPSSKSHSRAASLMMNMLPVGOK----- 679
 DB 902 --NLSPANISALATVGDLQREAVYMASNMANKTEALFYDLIREQREKLSGNDIY 959
 QY 680 -TTRHKFENKINITCP--NOVQPLRTYKNSMOGHTWLPQRDRA----- 721
 DB 960 YIARH-FEH-----VCPMPFAVMVPIIAAFSEBLQ-----LSSDPLILQSLDGRFLANN 1008
 QY 722 -----PSALHPTSVHALRPADIQVVALGDSLTLGANGIG 755
 DB 1009 VIFPFSMDLPNNAFQUTLKFTHLNNSTSLKVTNNHALKTL-LTSLAHGDKLR-----D 1062
 QY 756 SKPDDLPLVTITQYRGLSYSGAD-GSLENVITTLPLILBEPNNLGVAVGTODANDTNAP 814

```

Db 1063 SMKDVLLICISQLERVQLISAGVDINSLPDVSTTKPLKSLDKINRQSRGSSISLKHKSXF 1122
Qy 815 LNVQAVPAKAKEDLMSQVOTLMQKKDKDRNHFEDMKVITVLIGSLDCLCYCDSDNLYSA 874
Db 1123 QASSTHSTKS-----SSVEIVRE-----YSSRRVNAV-----DML-FSNTRNLGSE 1163
Qy 875 ANFDHLRNALDVLHREV-----PRV-LVNLVD-FLNPTIMROV-----FL 913
Db 1164 GIV-DFPKALIEVSWIEIEGSLSNRRLSLQCLVEISTYN--MRIMEMSSISL 1219
Qy 914 GNP-DKCPVQQAIVLNCVLTRENSQELARLEAFSRAVRSNRELVSGRYDTQDFPS 971
Db 1220 GTVFQVSCHENSITIAFALDSLRFQSMQFLIEBELS-----HFKFQDF- 1264
Qy 972 VLVQPF-----FQNLQPLVADGLPPTSPFADCIHPNCKFSQALATMTNMLEL--GSK 1026
Db 1265 --LQPFHAMENSODLKIKD-----LVLRCDQMIKARYONIRSGMRTIFHILAYASK 1315
Qy 1027 TETLDLAEPMPTCTPTQNEPFLRTPRNSNTYPIKPAIENWGSDFLCTEWKASNSVPTSV 1086
Db 1316 IENL-----LVLQCA----- 1325
Qy 1087 HQLRPADIKYVAALGDSLTTAV--GARPNNSSDLPTSMGLSWSIGDGNLETHTTLPN 1143
Db 1326 -----ISVSSLSGHEHISCVLTOGAYIDLISCI-TKPAKL-----NGNCKECLSCVD 1371
Qy 1144 ILKKEN-----PYLGFSTSTWGTAGLVAAGAR-- 1174
Db 1372 MLKLEHLLIKHLKMKKESVYSKKLEEWYLPPLSLFENELICE-ASDLEVRSKALVLP 1430
Qy 1175 -----ARDMPAQAMDLYER-----MKNSPDLNEKD-----WKLVTLLFGVN 1211
Db 1431 DCLYRHADDDEERWETVNSKALLISTISITNSQILYAKTBEVEMLTMTV-- 1486
Qy 1212 DLCHYCNPPRAHLATEVYQHIQALDILSEELPRAFVNVVEWELASLYOGQGGKAMLA 1271
Db 1487 -----BALKA-FIELDIKNLFERLHFLPRA-INTLE----- 1515
Qy 1272 AQNNCTGLRHSQS-----LEKQELKVVNNNLHGHSFSYWHQYT--QRED 1316
Db 1516 ---KCIQENSMISKVGLSCFSQFVLKKNQKQFKVDMD--EININSIQLDMLTPIELRD 1570
Qy 1317 FAVVQPFQNTLTPLNERG--DTDLTFESEDCHFSFDRGAB-----MATAIW 1363
Db 1571 PSLYPOVNSSSLDEVENSFRPRHEIRFNSQSVFKSKKHILKSIVVKTLLQMLMNCIW 1630
Qy 1364 -----NNMLEPVGK 1373
Db 1631 ELFHSNMLTNIPKR 1645

```

```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gccayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Chame M., Pfeiffer B.D.,
RA Anril J.F., Agbayani A., An H.-O., Andrews-Plambeck C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu I., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Bochan M.R., Bouck J., Brostein P., Broctier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobery C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Paclet J.M.,
RA Palazolo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier B., Spradling A.C., Stapleton M., Strong R., Sun B.,
RA Svirbas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhang L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [3]
RP FUNCTION.
RX MEDLINE=91092252; PubMed=2124970;
RA Keith F.J., Gay N.J.;
RT "The Drosophila membrane receptor Toll can function to promote
RT cellular adhesion."
RL EMBO J. 9:4299-4306(1990).
CC -1- FUNCTION: REQUIRED FOR DORSAL-VENTRAL EMBRYONIC POLARITY. MAY
CC -1- FUNCTION AS A MEMBRANE RECEPTOR. PROMOTES HETEROPHILIC CELLULAR
CC ADHESION.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.
CC -1- SIMILARITY: Contains 21 leucine-rich (LRR) repeats.
CC -1- SIMILARITY: Contains 1 TIR domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M19969; AAA28941.1; -
CC EMBL; AE003758; AAF56624.1; -
CC PIR; A29943; A29943.
CC HSP; O60603; IPIW.
CC FlyBase; FBgn0003717; TL.
CC GO; GO:0004888; F:transmembrane receptor activity; NAS.
CC GO; GO:0006963; P:antibacterial polypeptide induction; IMP.
CC GO; GO:0006966; P:antifungal humoral response (sensu Inverteb. .; IMP.

```

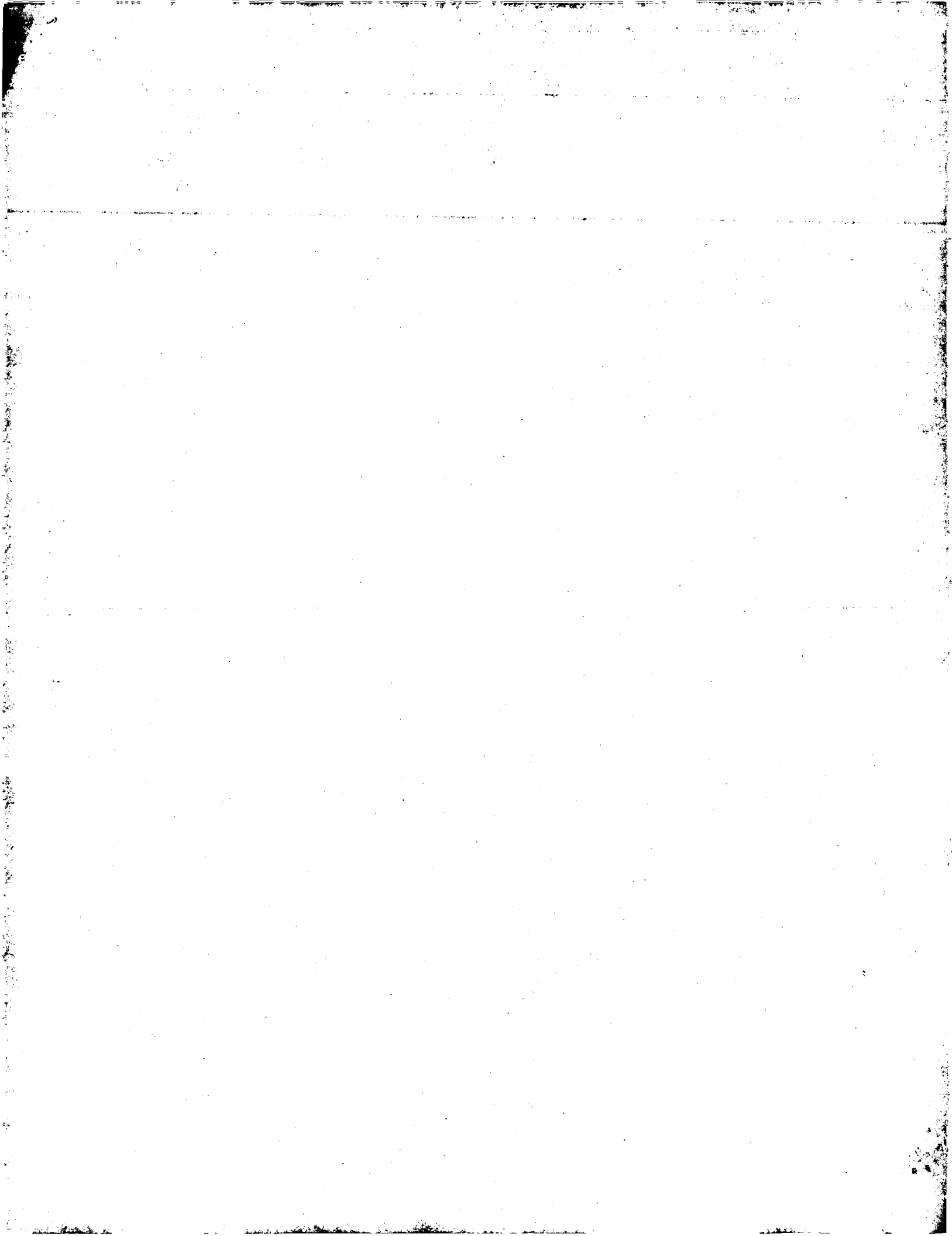
DR GO; GO:0006967; P:antifungal polypeptide induction; IMP.
 DR GO; GO:0009950; P:dorsal/ventral axis specification; NAS.
 DR GO; GO:0008063; P:rl receptor signaling pathway; NAS.
 DR InterPro: IPR004075; IL1_receptorl.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000483; LRR_Cterm.
 DR InterPro: IPR000372; LRR_Nterm.
 DR InterPro: IPR000157; LRR_Typ.
 DR Pfam: PF00560; LRR_11.
 DR Pfam: PF01463; LRRCT; 2.
 DR Pfam: PF01462; LRRNT; 1.
 DR Pfam: PF01582; TIR; 1.
 DR PRINTS: PRO0137; INTERKINRIF.
 DR PRINTS: PRO0019; LEURICHRPT.
 DR SMART: SM00369; LRR_Typ; 3.
 DR SMART: SM00082; LRRCT; 2.
 DR SMART: SM00013; LRRNT; 1.
 DR SMART: SM00255; TIR; 1.
 DR PROSITE: PS0104; TIR; 1.
 DR Developmental protein; Glycoprotein; Repeat;
 KM Cell adhesion; Leucine-rich repeat; Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 1097
 FT DOMAIN 28 807
 FT TRANSMEM 808 828
 FT DOMAIN 829 1097
 FT REPEAT 149 172
 FT REPEAT 174 195
 FT REPEAT 196 219
 FT REPEAT 221 243
 FT REPEAT 245 267
 FT REPEAT 268 291
 FT REPEAT 317 340
 FT REPEAT 341 364
 FT REPEAT 365 388
 FT REPEAT 390 412
 FT REPEAT 413 436
 FT REPEAT 438 459
 FT REPEAT 472 495
 FT REPEAT 496 521
 FT REPEAT 523 546
 FT REPEAT 550 574
 FT REPEAT 667 690
 FT REPEAT 692 713
 FT REPEAT 714 736
 FT REPEAT 785 807
 FT DOMAIN 857 996
 FT CARBOHYD 80 80
 FT CARBOHYD 140 140
 FT CARBOHYD 175 175
 FT CARBOHYD 235 235
 FT CARBOHYD 270 270
 FT CARBOHYD 275 275
 FT CARBOHYD 346 346
 FT CARBOHYD 391 391
 FT CARBOHYD 482 482
 FT CARBOHYD 508 508
 FT CARBOHYD 528 528
 FT CARBOHYD 654 654
 FT CARBOHYD 677 677
 FT CARBOHYD 702 702
 FT CARBOHYD 715 715
 FT CARBOHYD 730 730
 FT CARBOHYD 738 738
 SQ SEQUENCE 1097 AA; 124655 MM; D1BFC422453E3EABE CRC64;

Query Match 1.6%; Score 121.5; DB 1; Length 1097;
 Best Local Similarity 18.3%; Pred. No. 7.4;
 Matches 155; Conservative 109; Mismatches 315; Indels 269; Gaps 37;

668 LNMNMBPVGQKTRKFE-----NKININPNQVQPFRT 703

Db 228 LMSNQL-----HNLTKDFSGATSVLCIDIDHNGIEQLPHDVFAHLTNVTIDINLSANLFRS 283
 Qy 704 YKNSMGHGHTWP-----CNDRAPSAHPTSVAHARADQVAVALGDSLTAGIGISKPD 759
 Db 284 LPQGLFDHKKHLINEVRLMNNRVPLATLPSRLPANQP-ELQILRLRAE-----LQSLPG 335
 Qy 760 DLPDVTYQVGLSYSGDGSLE-----NTYTLPNILREPNRL 798
 Db 336 DLFHSTQLTNLSL---GDNLKITLPTLLEHQVNLISLDSNNRLTHLPDLSLPAFTNL 392
 Qy 799 TGYAVG---TSDANDTNAFLNQAVPGAAE----- 825
 Db 393 TDLRLIEDNLITGISGIFSNLGLVTVLVSRRRLRTIDSAFVSTGRLHLHDNDIDL 452
 Qy 826 -----DLMQVQ-----TLNQKMDDRVNHHEMKYITVLIGSDLCDCYCTDSN 870
 Db 453 QOPLDIDMLQTOINSPFGYHGLTLNLRNNSIIFVYNDMKYMTQLREIDL-SYNNISS 511
 Qy 871 L-YSAANFVDHLRNALDVHREVPRVL-----VNLVDFLNPITIMROVFLGNDPKCPVQAS 925
 Db 512 LGYEDLAFISQNRHLVNMTHNKIRIALPEDVHLGEGYNNLVHVDLNDP----- 562
 Qy 926 VLCNCVLTLENSOELARLEAFSAVSSMRBELVSGRYDQEDPSVVLQPFQNIQLPV 985
 Db 563 LVDCDITLFIQLVGVNHKPOYSRQFKLRTDRVCS-----QPN 601
 Qy 986 LADGLPTISFPAQDCIHENQKHSQALRALMTNMLEPLGSKTETDLRAEMPTIC--PTQ 1043
 Db 602 VLEGTFVRO-----LEPO-----TLICPLDPSDPRRKCPRGNCCHRTY 642
 Qy 1044 NEPELTPRNSVTVYIKRAIENWSGDFLCTEMKKSNSVPTSHQARPADIKVAL--- 1100
 Db 643 DIALVINCNSGNDTH--VRLPVLHKMQMLLEHLENNTLRLPSANTGEYSVLSHLA 700
 Qy 1101 GDSLTTA-VGARPNSSDLPTSMRGLSWSIGGDNLEHTTLNLIKKNPYLLG--ST 1157
 Db 701 GNNLISIDVDQPLTNLTHDISW-----NHLQMLNATVLSGLNRT 740
 Qy 1158 STWEGTAGLVAAG-----ARADMPAQAANDLYERMKNSPD--INLEKMKLVTLFI 1208
 Db 741 MKWR-----SVKLSGPNMCDCTAKPLLFTQDNFERIGDRNEMCMVNAEMPTRVEL-- 793
 Qy 1209 GYNDLCHYGENEBAHATEYVQHIOALDILSEELPRAVNVVEVNEIASLYOGGCKA 1268
 Db 794 STNDIC-----PAKGVFIALAVIALTGLAGFT-----AALYKFKQTEIK 835
 Qy 1269 MIAQNNCTCLRHSOSLEKOEKKVNMNLOHGISFSYWHQYTOREDAVAVVQPFONT 1328
 Db 836 IMLYAHNLLMLVTEEDLDKQ-----KFDAFISYSHD-----QSFIEDY 876
 Qy 1329 LPLPLNERGDTLTFSEDCFHFSQ---RGHAEMALAMNMMEPV--GRKT--TSNNFT 1380
 Db 877 LVQGLEHGQKQQL-----CVHERDMLVGH-----IPENIKRSVADSRTIIVLSQNF 926
 Qy 1381 HSR-AKLK 1387
 Db 927 KSEWALE 934

Search completed: January 6, 2004, 19:07:24
 Job time : 44 secs



GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2004, 14:46:04 ; Search time 44 Seconds

(without alignments)
3186.680 Million cell updates/sec

Title: US-10-054-691-2

Perfect score: 7766
Sequence: 1 MGRPGIFLELLILLGGT.....RCRRGRREDPWSRTVAL 1458

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: PIR 76:*
2: p1r1:*
3: p1r2:*
4: p1r3:*
5: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5987.5	77.1	1458	2	A45665 adult-specific brn
2	613.5	7.9	382	2	T20655 hypothetical prote
3	605.5	7.8	414	2	T26083 hypothetical prote
4	463.5	6.0	981	2	T16060 hypothetical prote
5	406.5	5.2	348	2	T24016 hypothetical prote
6	347	4.5	349	2	T24015 hypothetical prote
7	338	4.4	425	2	T21835 hypothetical prote
8	159.5	2.1	941	2	F97353 uncharacterized co
9	149	1.9	3655	2	T38084 TRAP-like protein
10	146.5	1.9	1755	2	F82618 chemotaxis-related
11	141.5	1.8	913	2	T15278 hypothetical prote
12	133	1.7	1708	2	F69189 protocorphyrin IX
13	131.5	1.7	1361	2	T30884 neural specific DN
14	130.5	1.7	891	2	E96590 hypothetical prote
15	130.5	1.7	1165	2	T15279 hypothetical prote
16	130.5	1.7	1545	2	T26589 hypothetical prote
17	129.5	1.7	1245	2	T42920 hypothetical prote
18	128	1.6	589	2	G86550 polysaccharide bio
19	127.5	1.6	1016	2	T30553 disease resistance
20	127	1.6	1280	2	T34357 hypothetical prote
21	126.5	1.6	1498	2	E86302 hypothetical prote
22	126.5	1.6	2471	2	T42977 large regement pro
23	126	1.6	1028	2	E51173 myosin I beta - bu
24	125	1.6	1178	2	S78475 mannosylphosphoryl
25	125	1.6	1650	2	S28721 mannosylphosphoryl
26	124	1.6	1946	2	AC2141 serine/threonine k
27	123.5	1.6	1230	2	B64664 outer membrane pro
28	123.5	1.6	1534	2	G91017 probable membrane
29	123.5	1.6	1534	2	A85862 hypothetical prote

30	123.5	1.6	2672	2	A48126 translation activa
31	122	1.6	1811	2	T39252 probable protein t
32	121.5	1.6	1097	2	A29943 Toll protein precu
33	121	1.6	4725	1	A44357 dynein heavy chain
34	120.5	1.6	836	2	D97182 extracellular neut
35	120.5	1.6	1570	2	T38792 probable protein t
36	120.5	1.6	2843	1	RBHUP adenomatous polyo
37	120	1.5	1237	2	D71850 probable outer mem
38	120	1.5	2140	2	P95074 serine proteinase,
39	120	1.5	15281	2	G41309 cyclosporin synthe
40	119.5	1.5	1163	2	E71481 probable swf/bnf h
41	119.5	1.5	1276	2	S11455 botulinum neurotox
42	119	1.5	1107	2	T25450 hypothetical prote
43	119	1.5	1116	2	T16112 hypothetical prote
44	119	1.5	2089	2	G85426 ATM-like protein l
45	119	1.5	3738	2	T05501 hypothetical prote

ALIGNMENTS

RESULT 1	
A45665	adult-specific brush border esterase/phospholipase (EC 3.-.-.-) precursor - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)	
C:Date: 03-May-1994 #sequence_revision 03-May-1994 #ext_change 08-Oct-1999	
C:Accession: A45665	
R:Boil, W.; Schmid-Chanda, T.; Semenza, G.; Mantel, N.	
J. Biol. Chem. 268, 12901-12911, 1993	
A:Title: Messenger RNAs expressed in intestine of adult but not baby rabbits. Isolation c	
A:Reference number: A45665; MID:93286138; PMID:8509424	
A:Accession: A45665	
A>Status: Preliminary	
A:Molecule type: mRNA	
A:Residues: 1-1458 <BOU>	
A:Cross-references: GB:Z12841; NID:91689; PIDN:CAW8303.1; PID:91690	
C:Keywords: hydrolase	
Query Match	77.1%; Score 5987.5; DB 2; Length 1458;
Best Local Similarity	77.6%; Pred. No. 0;
Matches 1130; Conservative 127; Mismatches 197; Indels 3; Gaps 3;	
QY	1 MGRPGIFLELLILLGGTPOIHTSPKSTIEGQWPEETLKNSPPCPNKLGVMPK 60
DB	1 MALMPSVFLGLPLGRGADQIQSSGNTLEGOLMPBSLTFPPCPKTLAEVPS 60
QY	61 SVSLKPSDIKFAAIGNLEIPDPDGTGLEKODWTE-RPOQVCMGVTVLSDIIRYFSP 119
DB	61 SVSLKPSDIKFAAIGNLEIPDPDGTGLEKODWTE-RPOQVCMGVTVLSDIIRYFSP 120
QY	120 SVPMPCVHTGKAVIPHDGAEDLWIOAQLVNMKNENLQDFQDWKLINVFSNMQCYL 179
DB	121 SALMPLC-DETRLVPRGAEDELWMOQTETLVSRBNPQDFEHDKLVNFFSNTSQCP 179
QY	180 CFSAQNGLAAGVDELMVLDYLOQEVPRARVNVLDSEVAEVSQYHGTWLSPPAERC 239
DB	180 CFSAQNGLAAGVDELMVLDYLOQEVPRARVNVLDSEVAEVSQYHGTWLSPPAERC 239
QY	240 NQSEETRLAKVVMQSVGEAANSLASRYESESFTVFPQFPFETPSLHSEDPRLQ 299
DB	240 RCLRETSQTLKVTQWQSYLEANDSLASRYESESFAVVFQFPFESSLSALLAEPPLQ 299
QY	300 DSTTLAMHLMNMMEPAGEKDEPLSVYGRPKMCPQSEPSYLSYNSNYLTRLQKPDK 359
DB	300 DSTTLALSLMNMMEPIGRKEPFSERKPKPCPTQSEPYLFTYNNSQLRVSPQPK 359
QY	360 LEVRGAEIRCDKPSDVTVPVSVRLKPADINVIGALGDSLTAGAGASTGNVLDVT 419
DB	360 LEVRGAEIRCDKPSDVTVPVSVRLKPADINVIGALGDSLTAGAGASTGNVLDVT 419
QY	420 QYRGLSMVSGDENIGTVTLTANILREFPNLSKGSVSGKETSPPAFNQAIVAGRAD 479

Dh	420	QYRGSLMSVGGDQINISTVTTTLANTILAREFNPSLQGRSVGTGHETTSSQAFNFQAVAGARADG	479
Qy	480	LPVQARRLVDMKQDTRTHFOEDMKITLFLFGNDLCDFCNDLVHYSFQNTDNI GKLD	539
Dh	480	LIPAQORVVALMKDTRINFQEDWKIIVFVIGNDLCDFCNDPVKYSFQNTDNI GTALD	539
Qy	540	ILHAEPAPFANLVTVLEIVNLRELJOEKYVCPEMILRSJCPCVLKCFDDNSTELATILE	599
Dh	540	ILHAEPAPFANLVKYLEISKRELJOQTKVSCPEMILRSJCPCVLKCFDDNSTELASILE	599
Qy	600	FNKKFOEKTHOLIESGRYDREDFTVVVQPFPEVNDMKTSBGLPDNSFFAPDCFHFSK	659
Dh	600	TIKEYQESTQOOLIDSGRYDREDFTVVVQPFPEKVMKPTODGLPDNSFFAPDCFHFSK	659
Qy	660	SHSRAASLMMNMLEPVQOKTTRHKFEKNKINTGNOVQPLFRKYSMOCHGTMLPORD	719
Dh	660	AHAHAASLMMNMLEPVQOKTTHNDEBEAVNITCNOVWPLSTYKNSVQGGTMLPORD	719
Qy	720	RAPALHPTSHALRPADIQVVAALGDSLITAGNGISKPDLPVTTQYRGSLYSAGDG	779
Dh	720	RSPASAPPTSHALRPADIQVVAALGDSLITAGIGISKPDNLSDOTTQYRGSLYSAGDG	779
Qy	780	SLEAVTTLLENILREBNRNLGTVANGTDANDTNAFLNOAVGAKAEDLMSQVOTLMQRMK	839
Dh	780	SLEAVTTLLENILRQFNSNLMGFAVGTGASGTAFAFNOAVGAKARDLMSQVOTLMQRMK	839
Qy	840	DDHRYNPFEDMKVITVLVIGSGLDCYCPDMSLYSAAVVDHRLNMLDYLHREVPVIVNL	899
Dh	840	DDHRYNPFEDMKVITVOJGASGLDCYCTDSNLYSAAVVDHRLNMLDYLHREVPVIVNL	899
Qy	900	VDLNPITMRQVFIGNPDKCEVQOASVLCNCVLTLENSOELIARLEAFSRAVRSMBELV	959
Dh	900	VDLNPITMRQVFIGNPDKCEVQOASALCNCVLSPRENSYELARLEALQAVOSSELRELV	959
Qy	960	GSGGYDPOEDSSVYLQPFQNIOLPYLADGLPDTSFAPDCLHPQKHSHQALRLMTMM	1019
Dh	960	ESGRDTRDESSVYLQPFPHSIQLEPVLODGRDITSFAPDCVHPQKHSHQSLRSLMTMM	1019
Qy	1020	LEPIGSKETETDLRAEMPICTPQONEPLRTPRNSNTYYPRIKPALENNGSDPLCTEMKRS	1079
Dh	1020	LEPIGSKETADLDTALITLCTPQONEPLRTPRNSDYTPSRPAVENNGSDPLCTAMNAS	1079
Qy	1080	NSVPTSVHQLRPADIKVVAALGDSLITTVAGARPNNSSDLPTSWRGLSNSIGDGNLETHT	1139
Dh	1080	RGVNPVSVAHELPGDIIKVVAALGDSLITTLTMGARPPSNSDPPMFWRGLSNSIGDGNLETHT	1139
Qy	1140	TLPNLIKKEFNYYLIGFSFTWEGTAGLVNAABGAPARMPAQAMDLVBERMNSPDIINEX	1199
Dh	1140	TLPNLIKKEFNYSILGFSFTLBEGLVGLNVAAGABARQAMPQARDLVBERMNSPDIIDEX	1199
Qy	1200	DMKLVTLFIVGNDLCYCENPENAHATEVYOHIOQALDILSEBELPRAEVNVVEWELASL	1259
Dh	1200	DMKLVTLFIVGNDLCHFCENPEGSGSEGEVYOHIOQALDIVELBELPRTFVNVVEWELAGL	1259
Qy	1260	YQGGGKGA-MLAAONNCTCLRHSGSSLEKQELKVVMMNLQHGISSBSYWHQYTOREFA	1318
Dh	1260	HODQGGRCATLLAQSHCTCKEYSSQSSVEMOELKKVMMNLQSGLSRYSYHQYOREFPA	1319
Qy	1319	VVVOPEFQNTLTPLNBERGDTLTFPSEDCPFHPSDGAHEMALMMNNMLPEPVGRYTTSSNN	1378
Dh	1320	VVVOPEFQNTLVPLNGRGTDLTFPSDDCFHPSSEGAHEMALMMNNMLPEPVGRYTTSSNN	1379
Qy	1379	FTHSRAKLKCSPPSPYLYTLRNSHLBPDOAEAEVLYMAVPVAGVGLVVGII GTVVM	1438
Dh	1380	FTYSRTKLKCSPPSPPYLYTLRNSRLLPDOAEADPTVLYMAVPVAGAGGLIGILAMVAG	1439
Qy	1439	RCRGRGREDPPMSLRT 1455	
Dh	1440	RGMRGCRREDPPLSLT 1456	

hypotheticalprotein F09C8.1 - *Caenorhabditis elegans*
C|Species: *Caenorhabditis elegans*
C|Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #ext_change 21-Jul-2000
C|Accession: T20655
R|McMurray, A.
submitted to the EMBL Data Library, November 1995
A|Reference number: Z19306
A|Accession: T20655
A|Status: preliminary; translated from GB/EMBL/DBJ
A|Molecule type: DNA
A|Residues: 1-382 <MW>
A|Cross-references: EMBL:Z68132; PIDN:CAA92221.1; GSPDB:GN00026; CESP:F09C8.1
A|Experimental source: clone F09C8
C|Genetics:
A|Gene: CESP:F09C8.1
A|Map position: X
A|Intons: 7/1; 71/3; 109/1; 164/3; 192/3; 245/2; 263/3; 303/3; 360/3
C|Superfamily: *Caenorhabditis elegans* hypothetical protein R0757.8

[illegible]

RESULT 3

hypothetical:protein W02B12.1 - *Caenorhabditis elegans*
 C.Species: *Caenorhabditis elegans*
 C.Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 01-Mar-2002
 C.Accession: T26083
 R.Swinburne, J.; Ainscough, R.
 submitted to the EMBL Data Library, October 1995
 A.Reference number: Z20149
 A.Accession: T26083
 A.Status: preliminary; translated from GB/EMBL/DBJ
 A.Molecule type: DNA
 A.Residues: 1-414 «WILL»
 A.Cross-references: EMBL:Z66521; PIDN:CAA91393.1; GSPDB:GNO0020; CESP:W02B12.1
 A.Experimental source: clone W02B12
 C.Genetics:
 A.Gene: CESP:W02B12.1
 A.Map position: 2
 A.Introns: 5/1; 61/3; 101/1; 156/3; 212/2; 273/2; 297/3; 332/3; 368/3
 C.Superfamily: *Caenorhabditis elegans* hypothetical protein R07B7.8

Query Match	7.8%	Pred. 605.5;	DB 2;	Length 414;
Best Local Similarity	36.8%	Pred. NO. 1.9e-33;		
Matches 141;	Conservative 57;	Mismatches 116;	Indels 67;	Gaps 10

RESULT 2
T20655

QY 376 SDVPSPSRHRLKPADINVLGALGDSLTLLGNGAGSTPGNVVLVLTQYKGLSMVSQGDENIG 43.5

Db 35 SKVPPSSASVSRPADIKYIGALGDSLTITANGAGAPGDPPLVITIQYKSLAFQIGGDSLD 94

QY 436 TVTTLANILREFNPSLKGFVSTGKETS- PNAFLNQAVAGRAEDLPVQARLVLDLKNKD 494

Db 95 EHITVANVLRKEPPLVLGASKIGSEBNWEVSHLMWVGAEISKDIIQARALVNTMHAI 154

QY 495 TRHHPEDMKIITLFFGSGNDLDFC-----NDLVHYSQNTTNDIGKALDILAEV- 546

Db 155 SEINVKEDWLVNIFIGANDICVYCEDPYFNSTALHGNATPEEKIIAVALQIDQNDERRT 214

QY 547 -----RAFVN-----LVTVLEIVNREL--VQEKVCVCPRIILR 578

Db 215 NEFKFLPSRI SVCKTPSSWREFRKNSNIYSRTIVSLTGMPNNRMLRKDIDKKYFEEGLATP 274

QY 579 SLCPCLAKEDNSTELATLIEFNKKFROEKTHO-----LIESGRYTRDEPTTV 626

Db 275 E-CDC-----ESNKQPTDDDIQGVCFGYMAEKDIQNTGQLEPNDODPTFV 318

QY 627 VQPFPEENV-DMPEKTSIGLEPDNSFFAPADCFHSSKSHSPASALANNNMLERPQCKTTRHKF 685

Db 319 VQPFPEGLIDDPYASGVVDMTFFAPADCFHSGVGHGNIQGNHINNTIIVQPVGKQTSVNL 378

QY 686 EN-KINITCPNQVQPLRTYKNS 707

Db 379 SDPSVGLHCPTNCPFPPTTKNS 401

```

RESULT 4
T16060
hypothetical protein F13H8.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16060
R:Ding, H.
submitted to the EMBL Data Library, July 1995
A:Description: The sequence of C. elegans cosmid F13H8.
A:Reference number: Z18455
A:Accession: T16060
A:Status: preliminary; translated from GB/EMBL/DBD
A:Molecule type: DNA
A:Residues: 1-981 <DIN>
A:Cross-references: EMBL:U23139; NID:G722370; PID:G722375; PIDN:AAC46681.1; CESP:F13H8.5
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP.F13H8.5
A:Initiators: 46/1, 124/1, 201/3, 255/2, 419/1, 452/1, 645/3, 666/3, 692/3, 719/3, 794/3, 8

```

[illegible]

```

Oy      653  CFFSSKSHSRASALMMNNMVEPVQKTKRHKFENKINI-----TCGNQVQFPLRYKNS 707
          |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      326  CFFLSIAIDIAIAKQIKWGLPEPIDQKTI-----INQLSGVDREPVCPVCEPIYRIQNS 381
          |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Oy      708  MGGHGHWLPCRD----RAPSAIHPISVAALPPADIQY 740
          |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      382  ENCE-----PSREMRFLRVPSAPASPGSLMPYIPIV 414
          |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

```

RESULT 5
T24016
Hypothetical protein R07B7.8 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #extc_change 20-Jun-2000
C:Accession: T24016
R:Harris, B.
submitted to the EMBL Data Library, July 1996
A:Reference number: Z19830
A:Accession: T24016
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-348 <WIL>
A:Cross-references: EMBL:Z75955; PIDN:CAB00118.1; GSPDB:GN00023; CESP:R07B7.8
A:Experimental source: clone R07B7
C:Genetics:
A:Gene: CESP:R07B7.8
A:Map position: 5
A:Intons: 4/1; 55/3; 91/1; 147/3; 223/2; 277/3; 335/3
A:Superfamily: *Caenorhabditis elegans* hypothetical protein R07B7.8

Query Match	5.2%	Score	406.5	DB	2	Length	348
Best Local Similarity	32.2%	Pred. No.	5.9e-20				
Matches	117	Conservative	66	Mismatches	133	Indels	47
						Gaps	11

```

OY 1060 IKALEIEMWGSDFLCTEWKASNSV-----PTSYHQLRPADIKVVAALGDSLTTVA GARPN 1113
Db 15 VKMALE-----KYNLSIDPHIEYDDVNMKPHHRIIVIGMGDSLTT--IGSRAE 61
OY 1114 NSSDLPTSMWGLSWSIGGDNLETHTTLPNLK---KPNPYLIGFSTSTWEG--TAGLNV 1168
Db 62 NIVGQRQRYPGNAFPTGMDEPVDRLHLLTVNVIIFRIAEKTNKGLFGSGSTGIDYGENTGILNV 121
OY 1169 AAEARARMDPAQAMDVERMKSSPDNLNEKDWLTLVTLFTGVND---LCHYCENPEAHLA 1225S
Db 122 AIGMKMSDDLIRQAKELVSRISKANKEININEMDWLTVSLWIGTVDVGTGLGYRLDEPIR--V 179
OY 1226 TEVYOHITQALDLISEELPRAPNVVVEWMEIASLYOOGGKGCMALAAQNNCTCRHSOSS 1285S
Db 180 DEYKSHIEKGLUYKENLPRTTIVSIYGMFPAQLIQEAQ-----SILKNGKRA 226
OY 1286 LEKOELKKVWNMLQHGSISSFYWHQYTOR--BDFAVVVOPF-FQNTLPLTNERGDTLTL 1341
Db 227 RLYENQKKLD-DLSDGYRNVSYDQNNHEHNSNDFLTVVQGPATRYDYSYRDEHGKKNPT 285
OY 1342 FFSFEDCHFSDRGHAEMALALMNNMLEPVGKRTTSNNFTHSRAKLKCPSPESPVLVTLRN 1401
Db 286 FYASDLTFHLTSKFGHAVLAKHYWNLTFEPVGKTRADLGCTKPKIYELNERNCILKTIVGN 345
OY 1402 SRL 1404
Db 346 SKM 348

RESULT 6
T24015
hypothetical protein R07B7.9 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
C/Accession: T24015
R/Harris, B.
submitted to the EMBL Data Library, July 1996
A/Reference number: Z19830
A/Accession: T24015

```


Db 342 NTLISGVHDNIGTQIOLSTGK-QTTLNMLEKMDWLFREFYGGDISPTPDK----- 388
 Qy 931 VTLRENSQELARLEAFSRAYSRSMRELVSGRYDQEDSFVVLQPFQNIQLPVLADGL 990
 Db 389 -----TDFTSLETRADNA-RSAMAMMIATRS-FGKKNINISGTAAPPKTINNLNLSNVN 438
 Qy 991 PD-----TSFFAPDCIHPNCKFHSQULARALMTNMLEPISGKTEPLDLRAEMPITCPTONE 1045
 Db 439 GDDVDTLALIVKNGSLTPQOKDYEAQLGVKAYATAFGIKINATADSA-----APRYVN 493
 Qy 1046 PLKLPFRNSNYTPPIKPAIENMGSDFLCTEW-----KASNSVPTSVHQLRADIKVVAALG 1101
 Db 494 -----SNEFTKQITITVPRK-----ATEYTLKLQASNGLSNS-----GISIVG 531
 Qy 1102 DSLITTVAGARPNNSSDLPTSRGLSMSIGDGNLET-HTLPLNLIKFNYPYLLGFSTISW 1160
 Db 532 KPTGTALVNPDSVTLNDRTTSTIDS--KGNVTTAPITSP-----MTYVYTYKVYDLG 582
 Qy 1161 EGTAGLNVAA--EGARARD-----MPAQA-----WDLVERMKSPDINLEKDM 1201
 Db 583 QAASGVITVASVNDGSGTDSSENFGLIIPANALISEVAGSGQFGYISTILN-----NIDTAS 638
 Qy 1202 KLVTLFIGNVNDLCHYCENPEAHPLATEYVQHIQQA-----LDI-----LSEELPRAFY 1248
 Db 639 SLITFLYGAG-ATYSIDIKSCTKVSDFQASQOSIFKMGVNDVQTQMSRLSDDDVKQFM 697
 Qy 1249 NVVEWELASLYOGGCKMMLAONNCTCLRHSSQSLERKELKKNWNLOHIGISFSSTY 1308
 Db 698 NY-----GKANIKNVITDITLTLNLSINIASIOKEKA-----TLSDMLPS-NYF 737
 Qy 1309 HOYTQREDAFVAVVQPFQONTLPLEN 1333
 Db 738 SQVSSD-----LQAWYNQTMDTLN 756

RESULT 9
 TRAP-like protein - fission yeast (Schizosaccharomyces pombe)
 C/Species: Schizosaccharomyces pombe
 C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C/Accession: T38084
 R/Gentles, S.; Odeall, C.; Churcher, C.M.; Barrall, B.G.; Rajandream, M.A.; Walsby, S.V.
 submitted to the EMBL Data Library, November 1995
 A/Reference number: Z21768
 A/Accession: T38084
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-3655 <GEN>
 A/Cross-references: EMBL:Z68136; PIDN:CAA92239.1; GSPDB:GN00066; SPDB:SPAC1F5.11C
 A/Experimental source: strain 972h-; contid c1F5
 C/Genetics:
 A/Genes: SPDB:SPAC1F5.11C
 A/Map position: 1
 A/Introns: 22/1

Query Match 1.9%; Score 149; DB 2; Length 3655;
 Best Local Similarity 20.0%; Pred. No. 1.5;
 Matches 167; Conservative 112; Mismatches 307; Indels 248; Gaps 40;
 Qy 672 MLEPVQOKTR--HKFKNKINITCPNOVQPLRTYKNSM-----OGHGWTMLPC 717
 Db 367 LARPLAFSTLADLLHVRDELNET---QIKSITMISTNMHDLTLSTGLQTMGARLLIM 423
 Qy 718 RDR---APSAHPSTVHALRPADIOVVAALGSLTAGNGSKPDDLPTVTOYRGISYS 774
 Db 424 VDRMISLPST-----PDALIFLLSIFDSFV-----NKFSELDNLDGPFKKKYE 467
 Qy 775 AGDGSLENTVTPNLIREFNRNLGVAVTGANDTNALNQAIVPQAKKEDLMSOVQTL 834
 Db 468 -----EELKETKSPTRSPRDLSPSTSVND-----GSFL-----FKNLMLFGIRAL 508
 Qy 835 MQMKODHRVNFHEWMKVIIVLLGGSDLCGYCTDSNLYSAANFVHDLRNALDVLHREVR 894

Db 509 MYGLR-----TCKSRCEIGCEQFSGFLTNIKXFEAVTF-----QKLFPEVGK 551
 Qy 895 VLVNVLDFLNPITMRQVFLNPKCPVOQASVLCNCLTLRENSQELARLEAFSRAIRS 954
 Db 552 GF---SYFAP--EGVYLETFFCC--EESLDRPALSTLPRKDEKDCLEAVATIF-- 599
 Qy 955 MRVLVSGRYDQEDSFVVLQPFQNIQLPVLADGLPD--TSFFAPDCIHPNQ---KPH 1008
 Db 600 -----IHLESIFLKVETN--LPTFPDQKKNLTTHIPOPFLSNESTSKFL 646
 Qy 1009 SOLARALMTNMLEPISGKTEPLDLRAEMPI-----TCPONEPFLR----- 1049
 Db 647 NILRFL-LSRIEELGSS-----DIRGQSVLRFLRSLFVTVSMFATENEVLRYVEEI 701
 Qy 1050 -----TPRNSNTYPIKPAIENMGSDFLCTEWKASNSVPTSVHQLRADIKVVAAL 1100
 Db 702 VKCMKLA PMSANSILNYLLRALFRIGGGRFESLYK--EVPPLHALLEAFNSLL-- 755
 Qy 1101 GDSLTTAVGARPNNSSDLPTS-----WRGLSMSIGDGNLET----- 1137
 Db 756 -----ISARTPEKDLFTTELCLTTPVRLSLLLPYSYMLRPLVMSLSKSOELVSGCLR 808
 Qy 1138 --HTLPLNLIKKE--NPLYLGFSSTW-----EGTAGLVNVAAGARARDM 1178
 Db 809 TPELCNDNLTPROFLDPIPMAYIEDLNNALMSHQPLRYVYNHSHTLKIKGLGGRKRL 868
 Qy 1179 PAQAMPLVERMKSPDINLEKMKVTLFLGVN--DLICGCENPEAHPLATEYVQHIQQA 1236
 Db 869 ---LDRVQSLKNSPSPN--NDPTLLLSIKGVQPOLH-----YQYVDEAV 910
 Qy 1237 DILSEELPRAFYVVE-----VMEASLYOGGCKMMLAONNCTCLRHSSQSLERKEL 1291
 Db 911 NLLSS--PSSDLEVKQOALFYVNCISKLYYKSDATNLSLSRCTADKSKSPDRRP 968
 Qy 1292 KKNWNLOHIGISSFSYHGYTOREDAFVAVVQ-----FPQNTLPLNERGDTLTFPSE- 1345
 Db 969 YSVIPSRMTGRSSFT---QISDSDDEIILASATYGLFAFTYDELRE-----EAYFLEK 1021
 Qy 1346 -----DCFHSD--RGHAEMALAMNNMLEPVGRKTTNNFTHSRAKLKC 1388
 Db 1022 LAVNVIVHDFYAFVDIYQGNHKSFTTNLQKEVI-----ISPHYAHCISEVVC 1069

RESULT 10
 F82618
 Chemotaxis-related protein kinase Xfr1952 [imported] - Xylella fastidiosa (strain 9asc)
 C/Species: Xylella fastidiosa
 C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
 C/Accession: F82618
 R/anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
 Nature 406, 151-157, 2000
 A/Title: The genome sequence of the plant pathogen Xylella fastidiosa.
 A/Reference number: AB2515; NCBI:20365717; PMID:10910347
 A/Note: for a complete list of authors see reference number A59328 below
 A/Accession: F82618
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-1755 <Sim>
 A/Cross-references: GB:AE004014; GB:AE003849; NID:G9107044; PIDN:AAF84754.1; GSPDB:GN001.
 A/Experimental source: strain 9asc
 R/Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
 Britones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre, H
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.
 submitted to Genbank, June 2000
 A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Fromm
 J.D.; Junqueira, M.L.; Kemper, B.L.; Kiteajima, J.P.; Krieger, J.E.; Kurame, E.E.; Laigri
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
 A/Authors: Martins, E.M.F.; Matsukuma, A.Y.; Nencik, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
 F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
 A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira
 M.; Teshato, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF1952

Query Match	1.9%;	Score 146.5;	DB 2;	Length 1755;
Best Local Similarity	18.7%;	Pred. No. 0.63;		
Matches 245;	Conservative 142;	Mismatches 457;	Indels 465;	Gaps 55

QY	193	YDELKGLVDLYQOEVPRAFVNLVDLSEV---	AEVRQYHGTG-LBPAPBCNCSSETRL	248				
Db	57	LDLALG---QARQELFEVENPVDTSRLISCDYH	HOVGOTLRMLETAPAMVAEEMERL	113				
QY	249	AKVVMQMSYOEAMNSLLASRSYEOESFTVFO---	PFYEYTPPSLHSEDPRLQDSTT	303				
Db	114	AKVMQ-----VDGVNDBREACAVLMRGTL	LLPDYLERLQNGHRDIP-----I	155				
QY	304	LAMHLMNFMMEPAGEKDEPLSVKGRPMKPC	SQESPYLFSTRNSYVLRLOQRPDKLEYR	363				
Db	156	VLLPLNLEIRARGE-----SRLENVLP	AFADP-----EVR	186				
QY	364	EGAEIRCDKOPS-----DTVP	TSVH-----RLKPA-----DIVVI	394				
Db	187	NATEYLDHANSLLTGCRQLDITASAKKE	LBELVKEALDYLRIGDVTDLRQITEL	246				
QY	395	GALGDLT-AGNG-AGSTPGNVLDTLYOGR	GSWSVGDENI--GTVTLIANIREFNS	450				
Db	247	SSVADTLMWGMGMSL-----VLQGR	DALLRYESEKEDAEMLLEINAGALLHV	300				
QY	451	LKGSFVGKETSPPNAFLNQAVAGRABDL	PVOARLVDMKNDTRIHFOEDWKITLFTI	510				
Db	301	LDNQVANIAGAEADAVULANA-----	ETRRTLDVLHEAISMETAARECFIAFI	349				
QY	511	GGNDLCPDCNDLVHNSPQNFNDICKAL	DLIDIAEVPRAFVNLVYLEIVNLAE	VOEKV	570			
Db	350	ESN-----WDQARLIKPEGLSEVG	ASLILSQASLEGEV-----RRTI	391				
QY	571	YCPRMILSLCPVLKFDNSTELATLLE	FNKKFOEKTQHLESGRYDREDFTVVO	PF	630			
Db	392	EC-ELITRKRPVAAQDLTDADAMAS	LEYELALR-----RRPERDILBET	TRNS	441			
QY	631	FENVD---MPKTSBGLPDNSFPAPDC	FHSSKSHSRASALMNNMLEPVQK	XTTHKEN	687			
Db	442	LEALSYPPIP-AEBGDSVGSQOST	DVLTFPSADP-----EALPEL	TVEDGTVPQEA	VSQ	495		
QY	668	KINITCPNOVQFELTKYNSMGHG	GNVL-----PCDRAP	SALHP	PS	729		
Db	496	SULTVEPSADIFDASLELVOP	IYTDLYEGDRSELPTESPMPDE	BPVIOISIP	VAA	PLS	555	
QY	730	-----AHALRPADIQVVAALG	SULTAGNG-----IGSKPDL	PPVTVYGL	SY	774		
Db	556	SGGFEDDVGIIDISIRDFL--	EELDERQGNLRLLTIWGA	PPDM--D	SLRQIR	RV	FHT	612
QY	775	AGDGSLENVTLPN-----	-----LIREFN	NLTGY	801			
Db	613	LKGSRLVCAITLGHFAWKIENTL	INRVLDSRAASPAAVALVERACE	VLPEMNAALR	SG	672		
QY	802	AVGTGDADDTNAFLNQAVGAKAED	MLSGVOTLMQKMDHVNHFEDK	VI	TVLIGSD	861		
Db	673	ASISDTLEATQALLDVLASGEBA	-----FYTSPSIAD	SAT-----LKL	PQ	LEG--	719	
QY	862	LCDYCTDSNLYSAANFVDHLR	NALDVLAREVRYVNLVD	PF-----NPTIM	ROY	911		
Db	720	-----GAPFVDSVLR--	ELEAEIANHLDITIDM	LTAAGOSY	SNESIL	RAV	766	
QY	912	FLGN-----PDKCPVQO-----	ASVLCN	CVL	TIRE	936		
Db	767	HTLNGAFAMTEWPELITQVWQ	SAETYIKRLTLKHQASVEGV	GLSAMATA	IA	DLTALRS	826	
QY	937	NSQELARLEAFERARSSREL	VSGRYTQODFSVLO	PPQNT-----QLP	VALDGL	PT	993	
Db	827	DS---PRIPSEV-SLVSC	SELV-----TTMPD	DGSGOQOYVER	SAEONL	KVPE	GTQV	877

QY	994	SEFAPDCIHPOKQSOLARLMTNMEPIGSKETJUDLBAEMITCTPTONPEPLATPRN	1053
Db	878	SL-----SNDVLTVDYDLSVNDVLEPLNAVVE-----PERIVDRPOE-----	917
QY	1054	SNVTYPIKPALENKGSDFLCTEAKSNSVPTSHQCLRPADIKVVALGDSLTUAVGARPN	1113
Db	918	-----VH0Y-----DIVAAS-----	929
QY	1114	NSDPLTSMRGLSWSIGDGNLEHTTLPNILKKNFYLLGFSTWEGTAGLVAAGA	1173
Db	930	-----MSDVVAADTNL-----	940
QY	1174	RADMPAOMDLVBRMKRSPDINL-----EKOMKLYTLPI-GVNDLCHYCENPEALHAT	1226
Db	941	---DVAAHQYDVLVG-NESPANELGLAFBEPABELVDLVEBSSDLDLDCNDNLATCL-994	
QY	1227	EYVOHIOALDILISEELPRAFVVNVEMELASLVOGSGCAMLAAONNCTCLRHOSOS-1285	
Db	995	-----NEAPQRELLVGLQR-DHHTLGG-ARWAGINALGDLGHSIESM	1036
QY	1286	-----LEKQELKXVMNLOHGHSISYSYHQ-YTGQEDRAPVVQOP	1323
Db	1037	LESVADYIMLNDRDMRLLEY-----SPDYHQMLQTRQHRHVMP	1078

RESULT 11

hypothetical protein R155.4 - *Caenorhabditis elegans*
 C1Species: *Caenorhabditis elegans*
 C1Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C1Accession: F15278
 C1Gensei: C.; Wamsley, P.; Kramer, J.
 submitted to the EMBL data library, May 1997
 A1Description: The sequence of *C. elegans* cosmid R155.

```

QY      646 NSFPAPDPCFHSSKSHSRASAL-----WN-----MLEPVQKTRHKE----- 686
Db      2 SEFF-----FPLAANNHNSFASFESILTTIMLNTCHGASPPVLNVPNSGLOEPHHDENYRGI 57
QY      687 --NKINITCPNQVQPLRTYKNSMQGHGWTMLPCRDRAPSLHPTSVALHPADIQVAAL 744
Db      58 GLNLAARRTKDELSTFTIGQFRSMAR-----LIT 85
QY      745 GDSLTLAAGNGIGSKRDD-----LPDVTYQRCISYGAGDG- 779
Db      86 GISLQGLAGIASIPPELISEFLNIGLTTTQVARIIDPVKLTEWQKVNTLSEFVSNDL 145
QY      780 -SLEN-----VTLLENILREFNRNLGYAVGTGD--ANDTNAFLNQAVP-----GAKAED 826
Db      146 KTIENHLYAMSIADKVKVGFK---GAYVVPDGDPTNLLDKFVSQYVTLDIISLAENLKD 202
QY      827 LMSQVQTL-----MQMKDHRVN----- 845
Db      203 FISSIELMKTIDITKTTKMEFYKIQSAITTISDIKAMFQFSGDSALNVIKTSLPFLSS 262
QY      846 ----FHEMWKIYTLIGSDLCDYCTDSNLYSANFADHLR-----NADVLVRE- 891
Db      263 VYSAVSEBEKIGKKTISIDSNLAEF--SKMSSSSNAAEELINKATDLYKLNQMWVYHHOK 320

```

```

QY 892 -----VPRVLVVOFLNPTIMROVFLGPRDCPCVQASVLCNCTLTRENSOELARLEAF 947
Db 321 HGNKVFKIPELID--GHTDITVFAEN-----SWLNDAYDELSCSEILGSLGV 368
QY 948 SRAVRSNRELVSGRYDTQ--EDFSVVLQPFONIQ-----LPVLADGLPDTSFFPAPDC 1000
Db 369 --ATDOAKETIMRIGSLATKLSQISPAISPLAGKVODIKIGSLPAESDK----- 415
QY 1001 IHRQKHSQLARLMTMTLEPLSGSKTETDLRAEMPTCTONEPPLTRPNSNRY-- 1058
Db 416 -----NLTFEKLITDM--ETLTSKVVALN--EM--VXTIND--LKKKSDALQRM 459
QY 1059 -PIKPAIENMSDELCTEMWKASNSVPTS HQLRPADIKVVALDSDLTAVAGAPNNSSD 1117
Db 460 STIDGVNNTNYDCKLQEVKASKDPVY-----NNLLVNNSSISFLNS 503
QY 1118 LPTSRGSLMSIGGDGNLEHTTLLPNIKKNPYLGFSTWEGTAGLVAAE----- 1171
Db 504 QPTM-----KKTAEISIAADALK--NPVULAIANOQIFEVINSILNTPPEIKPVIS 548
QY 1172 ----GARBDMPAQAMDVERKONSPOINIEKKMKVLTLFLGVNDLCHYCNPBAHLATE 1227
Db 549 AIKXFPFKTSVPVY--ITSVAKILPDI--KKDMKNLQTFVSKON--SNKTK----- 593
QY 1228 YVOHIQOALDILSEELPRAPVNVVEVWELASLYOGGCKCAMLAAONNCTCLRHSQSL 1287
Db 594 -----ESSVDIL--RELKNA---TVQLSVIGSVANGIFRMEQALGLTNDVSNMK--SFEAVV 643
QY 1288 KOELKKNVMNIQHGISFSFSTWHTQTOREDPRVVV-----OPFOCNTLTPLANER 1335
Db 644 KDEMSKYKLD-----GTDKKNMVLVILGLDELEKLSLDLNFKYSVKPKNST 690
QY 1336 GDTLTLFFSEDCPHFSDRGHAEMAI-----ALMNNNLBEPVGKRTT 1375
Db 691 NLADFPAGIFELAAAYKVGVPHNFMAIKASVLTLMBAANSNSLAGLLPIILN--LESIGLNF 749
QY 1376 SNNFTSRPAKLKCPSPESPVL--VTLRNSRL--PDQAEBAPEVL--YNAVPPVAAQVGLVVG 1431
Db 750 SSSFSKESKQSLKVLDTSPASLNTINNSRLKSTNAQCEPGLAKMHTIYIAGIPLIIT 809
QY 1432 I---IGTVVWRCRGR-----REDD--PMGLRT 1455
Db 810 IAVCVGVWTVWRRRSQRRRAAEAPAHNRDPPESRRS 846

```

RESULT 12
 F69189
 Protoporphyrin IX magnesium chelatase (EC 4.99.1.-) - Methanobacterium thermoautotrophicum
 C:Species: Methanobacterium thermoautotrophicum
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 24-Nov-1999
 C:Accession: F69189
 R:Smith, D.R.; Doncette-Stamm, L.A.; Delonghery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
 Qiu, D.; Spadafora, R.; Vicalaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jivanti, N.
 K.; S.; Church, G.M.; Daniels, C.D.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997
 A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func00
 A:Reference number: A69000; MUID:98037514; PMID:9371463
 A:Accession: F69189
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-1708 <MTH>
 A:Cross-references: GB:AE000847; GB:AE000666; NID:G2621756; PIDN:PAB85178.1; PID:G2621756
 A:Experimental source: strain Delta H
 C:Genetics:
 A:Gene: MTH673
 C:Superfamily: Methanobacterium thermoautotrophicum cobalamin biosynthesis protein N
 A:Keywords: Lyase

Query Match	1.7%;	Score 133;	DB 2;	Length 1708;
Best Local Similarity	20.7%;	Pred. No. 5;		
Matches 127;	Conservative 83;	Mismatches 202;	Indels 202;	Gaps 29;

```

0Y 593 ELATLIEKKFOCKT--HOLIEGRDPTREDPVTVVQPFENDMPETSGLDPSNFA 650
Db 181 D7SLIAYKQYPPBNTALHQMIDCALYYAAAGKTNLENQFTLAKTYNNRGIEFWNSWE 240
0Y 651 PDCEFFSSKSHSRASALMNMNMLEPVQCKTTRHKEFNKINITCENVOQFLR----- 702
Db 241 P-----ATVEPASPLSEBFLYRDQGRFKEDQYFTRYPDPDPKPTVAVLVSYGSTGEV 292
0Y 703 TYKSMOCHGTWLCRBRABALHPTSVHMLRPADIQVVALGSLTAAGNIGSKPDDL 762
Db 293 TYADAMQ-----QITDALVSRGLANVIPIG--TWSNV----- 323
0Y 763 DVTTQYRGLSYSAGDGSLENVTTLPNILEEF-----NRNLTYAVGTDANDTAFLN- 816
Db 324 -----LNGTQMOMIOTLCMPNOTYNTAIR-GIGVNTDLSILIG 364
0Y 817 QAVPGAKAEDMSOVOTLMQMKODHRVNFHEMKVITVLIGSSDLCDYCTDSNLYSAAN 876
Db 365 TSVSAN-----VVEVOJL-----DGNLIRSLKISTV-----QPVVYSA-- 400
0Y 877 FVHRLRNALDYLHREVPRLVNLVDPLNPTIMRQVPLGNPDKCPVOOASVLCNVLTR 936
Db 401 LVKELTDSNVVOYE-----ANPEKYV-KANVLIIDMLTFTG 437
0Y 937 NSQELARL-EAFSRAVRSKRELVSGRYDT-----QEDPSVY-----LOEPFONI 982
Db 438 STSGASVTKFRDSNPVLRAMITTSYRTIGOMIVSEBGFSSMWSYVQCAQEMOQ-Q 496
0Y 983 LPLVADGLPDTSFFAIPCDCIHFNQFHOQLARALMTNMLEPISGTETLIDLAENPITCPT 1042
Db 497 IEPLAIGVE-----IGSDPET-GAQMDITVITPE 525
0Y 1043 QNEBPF-----LRTPRNSN-----YTPPKPAIENMGSDPLCTEWRKASNSVPTSV 1066
Db 526 RIEKLVSRAFWIMILQTMANSDDKVALVYNYFP-PGKONIGASYL-----NVPEST 575
0Y 1087 HOLRRPADIKVVALGDSLTTAVGARPNNSSDLPTSW--RGLSWSIGDGNLETHTTLPN- 1143
Db 576 IEI-----LKRKAEGYS-----VGEIPODADALVEMMIKNGINVANMAPEGEKLANSSNA 627
0Y 1144 ILKKFNPYLLGFST 1157
Db 628 ILMPYEDLYAMENT 641

```

RESULT 13
T30884
neural specific DNA binding protein - African clawed frog
C|Species: *Xenopus laevis* (African clawed frog)
C|Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C|Accession: T30884
R|Marine, J.C.; Bellefroid, E.J.; Samain, H.P.; Martial, J.A.; Pieler, T.
submitted to the EMBL Data Library, December 1995
A|Reference number: Z20918
A|Accession: T30884
A|Status: preliminary; translated from GB/EMBL/DBD
A|Molecule type: mRNA
A|Residues: 1-1361 <MAR>
A|Cross-references: EMBL:U42462; NID:g150837; PID:g150838; PIDN:AAA98467.1
C|Genetics:
A|Note: Xg114

```

Query Match      1.7%; Score 131.5; DB 2; Length 1361;
Best Local Similarity 19.7%; Pred. No. 4.3;
Matches 249; Conservative 140; Mismatches 498; Indels 379; Gaps 58;

Oy 56 NMPKSVHS-----LKPSDIKF-----VAIGNLEIIPPPTGDTLEKQDWTERRPOOVG---- 103
    | | | | | : | : | : | | | | | | | | | | | | | | | | | | | | | | | |
Db 10 NSPTLSMISAARGISPAEVAHEHLKERGICYGLAPPPPGT-----TPREYCHQMA 59
    | | | | | : | : | : | | | | | | | | | | | | | | | | | | | | | | | |

Oy 104 -----MGVMTYLS-----DIIRYFSPVPMVCHTGRVIPHGDAG 139
    | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Db 60 FLASHASPYGELVQSAAGNTSLHLHYLTTPMDVSRFSSPRVYPRLL--SRKRALSTPLS 117
    | | | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```


QY 140 DLWIOAELVNMKENTOLDFOFDMKLIINVFESNASOCYLCPAQONGLAGVD----- 194
 DB 118 DASIDLOTMTIRTSNSL-----VAYINSSSSSSAASGYHLSGALSPASPF 165
 QY 195 -----ELMGVLDYLOQEVPRAPVNLVLSVAVSQYHG--TWLSPPAPPCNSEETT 246
 DB 166 PIHKPCSLAALAQ--QRSLSSFGHTPLHPSPTFASRQCALNSAPAPPSNNSAPDS 224
 QY 247 RLAVVNMQMSQVAMNSLL--ASSRYSQESFTVPF-QP----- 282
 DB 225 VLNKVSSESVAVSSTVNQVIRHRSKYKTEEBZSVRFQPPDHLTLDKEDLCKECKOPE 284
 QY 283 FFYEYETSLSEDRLODSTLAWHLNRMMEPAGEDE-----PLSVGH-- 327
 DB 285 HIYTNCHMOCCKSEPTQDQVHHINDHIIH--GEKKFPCRWQDCSRCKPFAQYML 342
 QY 328 -----GRMKPCSQ-----ESPFLSYRNSN-----YLTRL 353
 DB 343 VVHMRRTGEPKHCTEGECFKAYSRLNKLTHLSHTGKRPVYCDHEGCKAFNSASDR 402
 QY 354 QKPODKLEVEGAI--RCPDKDPDYPTVS-----HRLKPADINVGALGDS 400
 DB 403 AKQGNRHSNEKPYICVPGCTKRYTDPSSLRKHVKTVHGEBAHTVKKHRNDIIQK--PS 460
 QY 401 LTAGNGAGSTPGNVLDVLTQYRGLSMWSVGDE-----NIGVTTLANILREFPNLSKGFVS 456
 DB 461 LPKENG-----DNEASAKLSGREHSDVSXDQECLOTRTITKIDNMHMGSP--GGQS 512
 QY 457 GTCKETSPNAFLQAVAG-----GRAEDLPVQARLVLMKN-----D 494
 DB 513 SCSESPSPYGNNTNNDISGVDSLWQSLGDLFLLEFSPVVDSTVSMQSRSGPAPPET 572
 QY 495 TRIHFOE-----DWKITLFIGANDLCPCNDLVHSPONFTDNGKALDIILAE 544
 DB 573 QRHSAGTGAEREIKNERFILLIYEPNATCO--NRLPTISANGPDVGVSPSVLIN- 629
 QY 545 VPRAF--VNLVTVLEIVNLRELVOEKKVVCPRMILSLCVCVLFKPDNSTELATLLEPN 601
 DB 630 -PRAIELSMNDVYMMQNLNER-----DSTSTL----- 657
 QY 602 KKEQKETHOLIESGRYDTRDDFTVVVQPFENVDMPKTSB--GLPDNSFPAPDCFHRSK 659
 DB 658 -----SSAYTSRR--SSGISPPFSRSSESFQCGRLNNSSSADSDVPISTD 703
 QY 660 SHSRAALMMNLEPVGQKTRHKFENKINITC-----PNQVQPLRTYKXSMQ 709
 DB 704 ASRSSSEASQSHGLPNLNLTPAQHYRLKAKYAATGAPPTPLPNMDRIGLRNKLSLMD 763
 QY 710 GHGTWLP-----CRDRAPSA-LHPTSVALRPADIQVVAALDPSLAGIGSKP 758
 DB 764 GADPPLPFPFQOLPVPRRCSDGGNAGLTPWPHET--PGNNSRRAS--DPRARTAGIDDKP 820
 QY 759 DDLPDVTVQYRGLSYSGAGDGLSLENVTTL--PNILREFNRLTGYAVGTGDAN----- 809
 DB 821 --LPRFSRPH-----SNMSNMTLHPPLSBERNGSLCHYTOSDGLHHRVYSPR 867
 QY 810 ----DTNAFL-----NOAVPGAKAEMLM--SOVQTLMOQMKDHRVNFHEMKVITVLI 857
 DB 868 PPSISENVAMBAISCDADVPGGD--DDLMLPDDVQYIRSQNREAPENQLQTEYS----- 920
 QY 858 GGSGLDGYCTDSNLSYAA-----NFVDHLNALDVLHREVPV-----LVNLVDPLNPT 906
 DB 921 -----SPARNQSTTKSFHNTPEQPRAPGAYLSRNPALAECCGQTANMODNNMPV 972
 QY 907 IMRQVFLGNPD--KCPVQOASVLCNVLTLRENSQELARLEAFSRAVRSRRELVSGRY 964
 DB 973 QMNEVSSGTVDSVLPKQGFALA--GVLAVVQOQONQAYQSFQQAQMOGAHINMGQOE 1029
 QY 965 DTQEDFSVVLQPF-----FONIQLPVLADGLPPTSFPAPDCI--HPNOKHSOLA 1012
 DB 1030 SVQRNISVNGQRFENYLOQROQOMSOCCI-VSSDPIPOQRYVSQSQSMLSRAMQGGQOIS 1088

QY 1013 RALWTNNLEPLSGKXT-----ETLDRAMEPITCPTQ-----NEEFLRTPRNSNYTP 1059
 DB 1089 PSC--NNMVERPGVHTHAAPNTLNHQRALVHGAFTQGFANNSVNDGLHNP--NAYTVQF 1146
 QY 1060 IKPAIE 1065
 DB 1147 OKNGLE 1152
 RESULT 14
 E96590
 hypothetical protein T24C10.3 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: E96590
 R:Theologis, A.; Ecker, J.R.; Palm, C.U.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Hitzler, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, I.
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: E96590
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-891 <STO>
 A:Cross-references: GB:AEO05173; NID:g9657517; PIN:AA00872.1; GSPDB:GN00141
 A:Genetics:
 A:Gene: T24C10.3
 A:Map position: 1
 Query Match 1.7%; Score 130.5; DB 2; Length 891;
 Best Local Similarity 19.4%; Pred. No. 2.4;
 Matches 133; Conservative 92; Mismatches 290; Indels 171; Gaps 27;
 QY 156 LQDFQFDMKLIINVFESNASOCYLCPAQONGLAGVDL-----MGVL--DY 202
 DB 169 LNTGTERSVKASVNLMSSTLEVTLDIPYSDLGADKSSDDLDWEFKTASMGFTLGGSY 228
 QY 203 LQGEVPRAPVNLVLSVAVSQYHGTYMLSPAPPCNSEETTTL--AKVMQ----- 254
 DB 229 KOEBREKAVQNTADVSS-----GWSSPLNGTGPFPDPAKDAVLVAERENGDD 279
 QY 255 --WSYQEAWSLLASRYSEOE-----SETVVFOPEF--YETPSLRS--EDPRL 298
 DB 280 DPMVD--NGQWGFVALEAKPEKGRDLTNKESNGMGFGFEPVSKLETNNSFQSSVEKETKM 338
 QY 299 Q-----DSTLANHLMNRMMEPAGEDEPLSVHGHGPMKCPSESRYLPSYRN 346
 DB 339 ENGSISEFPGNDVNSGGTSVAFKQPSLEIGNEKEEK--EYOTGK--PKGVLPISFEDE 393
 QY 347 SNVLTIRQKPODKLEVBEGAEIRCPDKDPEDVPTS--VHRL--KPADIVIGALGSLT 402
 DB 394 KSSETDGLVHEDFVLASDEPVREKTKAPSPVTSISDLISRLSOVBKKAIVILENSAT 453
 QY 403 AGNGAGSTPGNVLDVLTQYRGLSMWSVGDENIQVTTLANILREFPNLSKGFVGTGET 462
 DB 454 ASNEVNGED-----SMEFGPKMPTLDSGIAGADDF-----STWFO 493
 QY 463 SPNAFLNOAVAGRAEDLPVQARLVLDLMDNDRIFHQ--EDMKITLFIGNDLDFC 519
 DB 494 GSPPALKMS-----DYTEVVDPEDDDS--WEFOGPTQVQKXMSRIGNGLMEX- 540
 QY 520 NDLVHVSPOQFTDNICKALDIILAEVPRAPVNL--VYTLLEIVNLRELVOEKKVCCPM 575
 DB 541 --KHSSVEVGN-----OSSVPNGFGEIHDXTVIRIENDVQDLHKIKIELYHI 589
 QY 576 ILRSL-----CPVLKFDNSTELATLIEFNKKFQKQKTHLIESGRYDTRDF 623

Db 590 ALVHELEKAKADSDSEVQKDEIEDLQNLNNDVLISGVNLESLQPGSSGMTELY 649
 Qy 624 TVVVOFFENV-----DMPKTSR-----GLPDSFPAPDC 653
 Db 650 KALQEKFRRLDEBDLTERLLSAEKDWKSTIELKATITLKIINLGSLEQSKASTW 709
 Qy 654 FHFS--KSHSRASALMNMMLPVGQKTRHKFEKNKINITCPNOVQPLRTYKNSMOGH 711
 Db 710 FEISTSCAOELRRAASIKWQVINKVDQEEITLSKPGKSVALSVEIYRVVKILRASTRLY 769
 Qy 712 GTWLPQRDRAPSAHPSTVALRPA 737
 Db 770 KPMI-----LLAPTSNVLA 787

RESULT 15

T15279

hypothetical protein R155.3 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C/Accession: T15279

R.Geisel C.J. Mansley, P.; Kramer, J.

submitted to the EMBL Data Library, May 1997

A/Description: The sequence of C. elegans cosmid R155.

A/Reference number: Z18321

A/Accession: T15279

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-1165 <GET>

A/Cross-references: EMBL:AF003390; NID:G208866; PID:G208870; PIDN:AA54273.1; GSPDB:GN

A/Experimental source: strain Bristol N2; clone R155

C/Genetics:

A/Map position: 3

A/Introns: 662/2; 741/1; 813/2; 941/2; 998/1; 1049/3

Query Match 1.7%; Score 130.5; DB 2; Length 1165;

Best Local Similarity 19.5%; Pred. No. 3.8;

Matches 185; Conservative 140; Mismatches 356; Indels 267; Gaps 44;

Qy 396 ALGDSITPAGNGASTPGNV--LDVLTQYKGLS--WSVGGDENIGTVTLTANILREFNPS 450
 Db 15 AYDEBLNTLKGAVDWGSIWTLIDELSNFLELSRLSMNDEMRLHLIKLKASVYTGTL 74
 Qy 451 LKGEFSGVGTGKSPN-----AFLNQAAGRA--EDLPVQAR---RLVDMKNDTRIHPQE 501
 Db 75 LVKYNYSQKEIVKSEVVYSLOKYYTSSIALSELPSKASNGDKIIDMLNS-----FEK 129
 Qy 502 DWKITP-----LFIGNDLDCFCNDL-----VHYSF 527
 Db 130 SIKSLTEDSPSICITLYELKSRHLRVSVALKTVSGFLNGSHDILHLFKDLEKMIKDAY 189
 Qy 528 QNETDNIGKALDILHA-----EVPRAFVNLVTLEIVMLRELYQEKVYCPRMILRLSLC 581
 Db 190 RDHSGMLANLNQKTFEASTRKVHOSFVSEAVMLLEVSK-----IRSKI 234
 Qy 582 PCVLKEDDDSTELATLIEPKKQEKTHQLI--ESGRYDREDFTVVVQPFENV--MP 637
 Db 235 P--KTSGLADKIATL-----PSKIHETLTPSVLPTSDK-----DKVDAKAP 275
 Qy 638 KTSSEGLPDSNFFAPDCFHFSSKSHSRASALMNMMLPVGQKTRHKFEKNKINITCPNOV 697
 Db 276 LMAQ-LPAQLKALEKLMNFSQKLPENNAKKLGNIIEIVNQ-----DGK-----NOL 321
 Qy 698 QPFLRTYKNS--MOGHGTWLPQRDR---APSAHPSTVALRPA--IQVVA--LGDSTL 748
 Db 322 ENTLKMKFSSTVDKEIYITLLPVKQDTLELKDNPIHETANIISLDKVEKFAQDMADPF 381
 Qy 749 TAGNGISGRPD--LPDVTQYKGLSYSGGDSLENTTLNLIREFNRNLTGYAVGT 805
 Db 382 TANNSIRQIESSLIIPIVIT--IRQLRLSNDQGLTTVNLNLPIMIKTOLS--SLTSY---- 435
 Qy 806 GDANDTNAFLNQAIVPGAEDLMSQVOTLMOQKKDHRVNFHEDMKVITVLLIGSDLCDY 865

Db 436 -----INSYKTKHAE-----TDALKDEKIGLHS--RVIGTATRGI ----- 470
 Qy 866 CTDSNLXSAANPFDHRLRNALDVLHREVPVVLVLPFLNPTIMROVPLGNPDKCPVQAS 925
 Db 471 ---SNQKLVDFKD-LADIGDLVKSEVEYKQQLND----- 502
 Qy 926 VLGNCVLTRENSQELARLEAFBRAYRSSMRLEVSGRYDTQDFSVVLQPFQNIQLPV 985
 Db 503 -----ENVANLKALAGIEGQLKTASGEIVG-----YKSVKPV 535
 Qy 986 LADGLPDTSPFAPDCIHP---NOKFHSQALRALMTNMLBPLGSKTETTLDRAMPITTCPT 1042
 Db 536 TSTSLPDYSTFPLDAKKVGINLKLPAIISMEKLNVLPIRSSRAT-----RSSST 586
 Qy 1043 QNEPFLTRNRSYTTPIPIPALENMSDPLCTEMKASNSVPTSVHQLRPADIRVVAALGD 1102
 Db 587 QLDGYLLSLQ-----SLSSISLDPT---KFKSGFDGTSLODLGVF--FAKYN 631
 Qy 1103 SLTTAVGARPNNSSDLPTSMRGLSWSIGDDGNLEHTTLPNILKKFPY---LLGFSST 1158
 Db 632 TLTKQKAAALAAQKSGGSS-----SASGSGGVFESNSTI--VLIVGCIIVICVILIPAS 685
 Qy 1159 TWEGTAGLVVAEGARA-RDMPAQAWDL-----VERMKNSPDINLEKDKLVTLFIGNV 1211
 Db 686 SWFYLHQDPAARKCPAIRELPWEVKKPVKKDVSKMKNKPERELEK-----TQDESIR 740
 Qy 1212 DLCHYENPEAHATBYV-----QHIOALDLISELPAPFNV 1250
 Db 741 DVPKLLELDMSAEKYVDRCRYEALQDMHKETCIPLSBELIKKFFGV 788

Search completed: January 6, 2004, 18:57:26
 Job time : 55 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2004, 18:33:20 ; Search time 370 Seconds
(without alignments)
790.635 Million cell updates/sec

Title: US-10-054-691-2
Perfect score: 7766
Sequence: 1 MGRPGIFILELLLLGQGT.....RCRRGRREDPPMSLRVAL 1458

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 733937 seqs, 200641211 residues

Total number of hits satisfying chosen parameters: 733937

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCRUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7766	100.0	1458	US-10-054-691-2	Sequence 2, Appl1
2	1392	17.9	270	US-10-094-749-2029	Sequence 2029, Ap
3	613.5	7.9	382	US-10-369-493-7025	Sequence 7025, Ap
4	605.5	7.8	414	US-10-369-493-5575	Sequence 5575, Ap
5	560	7.2	148	US-09-764-891-4078	Sequence 4078, Ap
6	463.5	6.0	981	US-10-369-493-5416	Sequence 5416, Ap
7	406.5	5.2	348	US-10-369-493-6523	Sequence 6523, Ap
8	347	4.5	349	US-10-369-493-6522	Sequence 6522, Ap
9	338	4.4	425	US-10-369-493-5161	Sequence 5161, Ap
10	141.5	1.8	799	US-10-104-047-2335	Sequence 2335, Ap
11	136	1.8	300	US-10-156-761-13112	Sequence 13112, A
12	133	1.7	1708	US-10-369-493-1156	Sequence 1156, Ap
13	129.5	1.7	1633	US-10-359-012-4	Sequence 4, Appl1
14	128.5	1.7	2765	US-09-808-602-84	Sequence 84, Appl
15	128.5	1.7	2765	US-09-800-198-72	Sequence 72, Appl

16	127	1.6	787	15	US-10-128-714-8587	Sequence 8587, Ap
17	124	1.6	2828	10	US-09-905-129-21	Sequence 21, Appl
18	124	1.6	2828	10	US-09-991-630-21	Sequence 21, Appl
19	124	1.6	2828	12	US-10-301-822-49	Sequence 49, Appl
20	124	1.6	2828	12	US-10-032-189-126	Sequence 126, App
21	124	1.6	2828	12	US-10-295-027-58	Sequence 58, Appl
22	124	1.6	2828	12	US-10-295-027-1175	Sequence 1175, Ap
23	124	1.6	2828	15	US-10-176-847-54	Sequence 54, Appl
24	124	1.6	2828	15	US-10-177-293-110	Sequence 110, App
25	123.5	1.6	1230	10	US-09-881-752A-150	Sequence 150, App
26	123	1.6	1335	12	US-10-369-493-11196	Sequence 11196, A
27	122.5	1.6	2764	10	US-09-808-602-80	Sequence 80, Appl
28	122.5	1.6	2764	11	US-09-800-198-68	Sequence 68, Appl
29	121.5	1.6	1276	12	US-10-354-774-66	Sequence 66, Appl
30	121.5	1.6	1276	12	US-10-271-012-66	Sequence 66, Appl
31	121.5	1.6	2597	10	US-09-905-129-2	Sequence 2, Appl1
32	121.5	1.6	2597	10	US-09-905-129-13	Sequence 10, Appl
33	121.5	1.6	2597	10	US-09-905-129-13	Sequence 13, Appl
34	121.5	1.6	2597	10	US-09-991-630-10	Sequence 10, Appl
35	121.5	1.6	2597	10	US-09-991-630-13	Sequence 13, Appl
36	121.5	1.6	2597	10	US-09-991-630-13	Sequence 13, Appl
37	120.5	1.6	2843	10	US-09-987-482-1	Sequence 1, Appl1
38	120	1.5	816	12	US-10-369-493-3409	Sequence 3409, Ap
39	119.5	1.5	1192	9	US-09-758-140-6	Sequence 6, Appl1
40	119.5	1.5	1192	9	US-09-752-599A-6	Sequence 71, Appl
41	119.5	1.5	1192	15	US-10-060-036-71	Sequence 28, Appl
42	119.5	1.5	2119	12	US-09-769-744A-28	Sequence 32, Appl
43	119.5	1.5	2843	8	US-08-681-219-32	Sequence 30, Appl
44	119.5	1.5	2843	12	US-10-092-138-30	Sequence 30, Appl
45	119.5	1.5	2843	12	US-09-230-111C-30	Sequence 30, Appl

ALIGNMENTS

RESULT 1
US-10-054-691-2
; Sequence 2, Application US/10054691
; Publication No. US20020115846A1
; GENERAL INFORMATION:
; APPLICANT: Yu, Xuanchuan
; APPLICANT: Mirand, Maricar
; TITLE OF INVENTION: No. US20020115846A1 Human Lipase and Polynucleotides Encoding ti
; FILE REFERENCE: LEX-0303-USA
; CURRENT APPLICATION NUMBER: US/10/054,691
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: US 60/264,049
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1458
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-054-691-2

Query Match 100.0%; Score 7766; DB 14; Length 1458;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1458; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGRPGIFILELLLLGQGTPOHTSPKSTLFGOLMPETLKNSPPCPNPKLGVMPSK 60
Db 1 MGRPGIFILELLLLGQGTPOHTSPKSTLFGOLMPETLKNSPPCPNPKLGVMPSK 60
QY 61 SVHSLKPSDIKFAAGNLEIPDPDGTGLDEKQDWTERRPOQVCMGVTLSIDIRFSPS 120
Db 61 SVHSLKPSDIKFAAGNLEIPDPDGTGLDEKQDWTERRPOQVCMGVTLSIDIRFSPS 120
QY 121 VPMVCHTKRKVI PHDGAEDLWIOAGELVRNKKENIQDFOFDWKLINVFNSAOCYLC 180
Db 121 VPMVCHTKRKVI PHDGAEDLWIOAGELVRNKKENIQDFOFDWKLINVFNSAOCYLC 180

QY 181 PSAOONGLAAGVDELMGVDLYLOEVPRAFNVLVDLSEVAEVSROYHGTWLSPAPEPCN 240
 DB 181 PSAOONGLAAGVDELMGVDLYLOEVPRAFNVLVDLSEVAEVSROYHGTWLSPAPEPCN 240
 QY 241 CSEETRLAKVWQWMSYOEAANSLASSRYSSEGESFTVVYQPFYEYETPSLSHSEBRLD 300
 DB 241 CSEETRLAKVWQWMSYOEAANSLASSRYSSEGESFTVVYQPFYEYETPSLSHSEBRLD 300
 QY 301 STTLAMHLMNRMEPEGEKDEPLSVYHGRPMKCPSESPILFSYRNSNTLRLQKQODLU 360
 DB 301 STTLAMHLMNRMEPEGEKDEPLSVYHGRPMKCPSESPILFSYRNSNTLRLQKQODLU 360
 QY 361 EYREGAEICRCPDSDPTVPTSVHRLKPADINVIGALGDSLTAGNAGAGSTPGVNLVLNQ 420
 DB 361 EYREGAEICRCPDSDPTVPTSVHRLKPADINVIGALGDSLTAGNAGAGSTPGVNLVLNQ 420
 QY 421 YRGLSNWSVGDEINIGTVTLTANILREFPNLSKGFSGTGETSPNAFLNQAVAGRAEDL 480
 DB 421 YRGLSNWSVGDEINIGTVTLTANILREFPNLSKGFSGTGETSPNAFLNQAVAGRAEDL 480
 QY 481 PYQARLVLMKMDTRIHFQEDMKITTLPIGNDLDCDFCNDLVHSGPQNTDNGKALDI 540
 DB 481 PYQARLVLMKMDTRIHFQEDMKITTLPIGNDLDCDFCNDLVHSGPQNTDNGKALDI 540
 QY 541 LHAEPRAFNVLTVLEIVNLRELVOEKKYCPRMILRSICPCLKFDNSTELATLIEF 600
 DB 541 LHAEPRAFNVLTVLEIVNLRELVOEKKYCPRMILRSICPCLKFDNSTELATLIEF 600
 QY 601 NKKFOEKTQOLISGRYDTRDEFTVVYQPFENVDMPTSEGLPDNSFPAADCFFHSSKS 660
 DB 601 NKKFOEKTQOLISGRYDTRDEFTVVYQPFENVDMPTSEGLPDNSFPAADCFFHSSKS 660
 QY 661 HSRAAALNMNMLEPGQKTRHFKFENKINITCPNOVOPRLTYKNSMOGHGWLRCRDR 720
 DB 661 HSRAAALNMNMLEPGQKTRHFKFENKINITCPNOVOPRLTYKNSMOGHGWLRCRDR 720
 QY 721 APSALHPTSVAHLPADIQVVAALGDSLTAGNGISGKPDPLPVYTYOYRSLYSAGDGS 780
 DB 721 APSALHPTSVAHLPADIQVVAALGDSLTAGNGISGKPDPLPVYTYOYRSLYSAGDGS 780
 QY 781 LENVTLLPNILREFPNRLTGAVGTGDANDTNAFLNQAVGAKAEDLMSQVOTLMQMKD 840
 DB 781 LENVTLLPNILREFPNRLTGAVGTGDANDTNAFLNQAVGAKAEDLMSQVOTLMQMKD 840
 QY 841 DHRVNHEDMKITTVLIGSGLDCDYCTDSNLYSAANVVDLRLAALDYLHREVRVLVNLV 900
 DB 841 DHRVNHEDMKITTVLIGSGLDCDYCTDSNLYSAANVVDLRLAALDYLHREVRVLVNLV 900
 QY 901 DFLNPTIMROVFLGNBPKCPVOQASVLCNCVLTILRENSQELARLEAFSRAYSRSMRELVG 960
 DB 901 DFLNPTIMROVFLGNBPKCPVOQASVLCNCVLTILRENSQELARLEAFSRAYSRSMRELVG 960
 QY 961 SGRYDQOEDSVVLQOPEFNONIQLPVLADGLPDTSFPAADCIPHNOKEHSQLARALWTNML 1020
 DB 961 SGRYDQOEDSVVLQOPEFNONIQLPVLADGLPDTSFPAADCIPHNOKEHSQLARALWTNML 1020
 QY 1021 EPLGSKTETDIAEAEMPTCTPTONEPRLTRPRNSNTYYPILKPALENNGSFPLCTEWAASN 1080
 DB 1021 EPLGSKTETDIAEAEMPTCTPTONEPRLTRPRNSNTYYPILKPALENNGSFPLCTEWAASN 1080
 QY 1081 SVFTSVHQLRPADIKVVAALGDSLTTAVGARPNNSSDLPSTSMKGLSNISIGDGNLEHTHT 1140
 DB 1081 SVFTSVHQLRPADIKVVAALGDSLTTAVGARPNNSSDLPSTSMKGLSNISIGDGNLEHTHT 1140
 QY 1141 LPLNLIKFNBYLLGFSTSTWEGTAGLVNVAAGARARDMPAQAAMDVERMKNSPDINLEKD 1200
 DB 1141 LPLNLIKFNBYLLGFSTSTWEGTAGLVNVAAGARARDMPAQAAMDVERMKNSPDINLEKD 1200
 QY 1201 WKLVTLFIGNVDLCHYENPEBAHLATRYVOHIOQALDILSEBELPRAVNVVEWELASLY 1260
 DB 1201 WKLVTLFIGNVDLCHYENPEBAHLATRYVOHIOQALDILSEBELPRAVNVVEWELASLY 1260
 QY 1261 QGGGKCAMLAQNNCTCLRHSSQSLKQELKKVNMNLOHGISFSYWHQYTORDEPAVY 1320

DB 1261 QGGGKCAMLAQNNCTCLRHSSQSLKQELKKVNMNLOHGISFSYWHQYTORDEPAVY 1320
 QY 1321 VOPFFONTLTPLENGDDTLTFSSDCFFHSPRGAEBAIALAMNMLEVGKRTSSNNFT 1380
 DB 1321 VOPFFONTLTPLENGDDTLTFSSDCFFHSPRGAEBAIALAMNMLEVGKRTSSNNFT 1380
 QY 1381 HSRALKKCPSPESPLYTLTLRNSRLPDQAEAPVLYMAVPAAGVGLVVGIGTVWRC 1440
 DB 1381 HSRALKKCPSPESPLYTLTLRNSRLPDQAEAPVLYMAVPAAGVGLVVGIGTVWRC 1440
 QY 1441 RRGGRREDPPMSLRVAL 1458
 DB 1441 RRGGRREDPPMSLRVAL 1458
 RESULT 2
 US-10-094-749-2029
 ; Sequence 2029, Application US/10094749
 ; Publication No. US20030219741A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ISOGAI, TAKAO
 ; APPLICANT: SUGIYAMA, TOMOYASU
 ; APPLICANT: OTSUKI, TETSUJI
 ; APPLICANT: WAKAMATSU, AI
 ; APPLICANT: SATO, HIROYUKI
 ; APPLICANT: ISHII, SHIZUKO
 ; APPLICANT: YAMAMOTO, JUN-ICHI
 ; APPLICANT: ISONO, YUUKO
 ; APPLICANT: HIO, YURI
 ; APPLICANT: OTSUKA, KAORU
 ; APPLICANT: NAGAI, KEIICHI
 ; APPLICANT: IRIE, RYOTARO
 ; APPLICANT: TAMECHIKA, ICHIRO
 ; APPLICANT: SEKI, NAOHICO
 ; APPLICANT: YOSHIKAWA, TSUTOMU
 ; APPLICANT: OTSUKA, MOTOYUKI
 ; APPLICANT: NAGAHARI, KENJI
 ; APPLICANT: MASUHO, YASUHIKO
 ; TITLE OF INVENTION: NOVEL FULL-LENGTH cDNA
 ; FILE REFERENCE: 084335/0160
 ; CURRENT APPLICATION NUMBER: US/10/094,749
 ; CURRENT FILING DATE: 2002-03-12
 ; PRIOR APPLICATION NUMBER: 60/350,435
 ; PRIOR FILING DATE: 2002-01-24
 ; PRIOR APPLICATION NUMBER: JP 2001-328381
 ; PRIOR FILING DATE: 2001-09-14
 ; NUMBER OF SEQ ID NOS: 3181
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2029
 ; LENGTH: 270
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-094-749-2029
 Query Match 17.9%; Score 1392; DB 12; Length 270;
 Best Local Similarity 100.0%; Pred. No. 2.4e-119;
 Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 181 NMLEPVGQKTRHKNKINITECNQVPLRTYKNSMOGHGTLWPCDRAPASALHPTSV 240
 QY 731 HALRPADIQVVAALGDSLT 749
 Db 241 HALRPADIQVVAALGDSLT 259

RESULT 3

US-10-369-493-7025
 ; Sequence 7025, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; PRIOR FILING DATE: 2002-02-21
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 7025
 ; LENGTH: 382
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 ; US-10-369-493-7025

Query Match 7.9%; Score 613.5; DB 12; Length 382;
 Best Local Similarity 38.9%; Pred. No. 3e-47;
 Matches 133; Conservative 60; Mismatches 111; Indels 47; Gaps 10;
 QY 376 SDTVPSTVRLKPADINIVGALGDSLTAG-----NGASTPGNVLDVLTQYRGLSWS 427
 Db 37 SKVPTNVNSVRPADIKLIMAGDSLTGVGLFSSQANAGAE--DVAVALVQYRGLAFO 94
 QY 428 VGGDENIGVTTLANTIRENPISLKGKSVGTGKETSNN---AFLNQAVALGGRADLPVQ 483
 Db 95 AGGDKTLEEHVTLPIILKKNPDPFGYSNGIG--SNVWEIARLNVAMFGANAKDIPGQ 151
 QY 484 ARRLVDMKNDTR-IHQEDMKIITLFIGNDLCPDNLDVHYSPOFTDNIGKALDILH 542
 Db 152 AKQLVQLQCHTEVNNKEDMKLNTIFGANDICGRKVEDSPYCAODIQOAVQIY 211
 QY 543 AEVPRAFVNLVTLVLEIYNRELVOEKKVYCPMILRSICPVLFKFDNSTELATLIEFNK 602
 Db 212 DNVPRVIVSLTGMLEMLAQT-DTGMWFCQR-LHDEGCG-----ESNK 254
 QY 603 KPE-----KTHQLESGRVDRDFVTVVQPFENVDMPKTSGLPDNSFPAP 651
 Db 255 NFDADIRQACDYNNKYEKQIENDGFEEKNDFTYVVOQPMFOPTLIPMENGKFTOKFFAP 314
 QY 652 DCFHSSKSHSRASALNMNMLEPVGQKTRHKNKINITECNQVPLRTYKNSMOGHGTLW 707
 Db 315 DCFHSGMGAHALVSTILMNNILQDPVGSKSTVSNMSPVLQTLACDAACPRIKIPKNS 371

RESULT 4

US-10-369-493-5575
 ; Sequence 5575, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 5575
 ; LENGTH: 414
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 ; US-10-369-493-5575

Query Match 7.8%; Score 605.5; DB 12; Length 414;
 Best Local Similarity 36.8%; Pred. No. 1.9e-46;
 Matches 141; Conservative 57; Mismatches 118; Indels 67; Gaps 10;
 QY 376 SDTVPSTVRLKPADINIVGALGDSLTAGAGSTGQNVLDVLTQYRGLSWSVGGDENIG 435
 Db 35 SKVPTNVNSVRPADIKVIGALGDSLTAAAGACQDPLAVILOVYRGLAFOIGDKSLD 94
 QY 436 TVTTLNIIREEPNSLKGKSVGTGKETS--PNAPLNQAVALGGRADLPVQARRVLDLMKND 494
 Db 95 EHTTVANVLRKFPVTLVGSAGKIGSENWVESHNLNGVPAESKDIIIGQKALVNTMHAH 154
 QY 495 TRIHQEDMKIITLFIGNDLCPFC-----NDLVHYSPOFTDNIGKALDILHAEVP--- 546
 Db 155 SEINVEDMKLVNIFIGANDICVYCEDPYFNSTALHGNATFEKNIIAAVOILQDNLPRT 214
 QY 547 -----RAFYN-----LVTLVLEIYNREL--YOEKKVYCPMILR 578
 Db 215 NFKKLFBSRISVCKTFSSWREFKNSNIYGRITVSLTGMFMRMRLRKIDKKYCEGLHTE 274
 QY 579 SLCPCLKDDNSTELATLIEFNKFOEKTHQ-----LIESGRVDRDFTV 626
 Db 275 B-CDG-----ESNKQPTDDIDQVCTCGYMAEKDIONTGLFNNKDDPTFV 318
 QY 627 VQPFENV-DMPKTSGLPDNSFPAPDCFFSSKSHSRASALNMNMLEPVGQKTRHKNK 685
 Db 319 VQPFENVGIDPRPASGVVDMTFAPDCFFHSGHGNIGMILMNTIVQVSGKQTSVNL 378
 QY 686 EN-KINITECNQVPLRTYKNS 707
 Db 379 SDPSVGLHCPSTWCPEFPPTKNS 401

RESULT 5

US-09-764-891-4078
 ; Sequence 4078, Application US/09764891
 ; Publication No. US20030077808A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PC006
 ; CURRENT APPLICATION NUMBER: US/09/764,891
 ; PRIOR FILING DATE: 2001-01-17
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 10231
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 4078
 ; LENGTH: 148
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-764-891-4078

Query Match 7.2%; Score 560; DB 11; Length 148;
 Best Local Similarity 77.9%; Pred. No. 4.6e-43;
 Matches 113; Conservative 11; Mismatches 17; Indels 4; Gaps 3;
 QY 1013 PALMTMLEPLSGSKTTLRLAEMPTTCPTONEPPLRTFRNSYTYPIKPALENNSSDL 1072
 Db 4 QSLMTMLEPLSGSKTTLRLAEMPTTCPTONEPPLRTFRNSYTYPIKPALENNSSDL 63
 QY 1073 CTEWKAHSVPTSVQOLRPADIKVVAALGDSLTAAAGAPNNSSDLPTSMRG-LWSIGG 1131

Db 64 CTEWKASVPTSVHQLRPADIKVVALGDSLTTAVGARNNSSEPHILEGTLGALLEG 123
 QY 1132 DGNL-ETHTLTPN--LKKNPYLL 1153
 Db 124 MGNIGDSHTTAQHSSEVQPLPMLL 148

RESULT 6 US-10-369-493-5416

; Sequence 5416, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 5416
 ; LENGTH: 981
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 ; US-10-369-493-5416

Query Match 6.0%; Score 463.5; DB 12; Length 981;

Best Local Similarity 29.7%; Pred. No. 1.2e-32; Indels 39; Gaps 14;
 Matches 118; Conservative 71; Mismatches 169;

QY 358 DKLEVRREG-AEIRCPDPSDPTVPTSVHRLKPADINIVGALGDSLTAAGAGSTPGNVLD 416
 Db 43 DSEMEYMPGNPNFNP--QPHKXSESVAQLHPSQIGVIALGDSVVAQAQKS--SILD 98
 QY 417 VLTQYRLASVSGDENVIGVTTTLANTLREPNBPKGFSVGTGKETSPPNAFLNQAIVGGR 476
 Db 99 LFDQFPVSFVTGDDVLTNEQATFINIFRKAPRIKQ---GSSDPVQKFDYDFNNAIIGSF 155
 QY 477 AEDLPVQARLVDMKMDTRTHPEQDWKIITLFFGNDLQPCNDLVHVSFONPTDINGK 536
 Db 156 SSELPEDALKLASTLTKKLGHETSDTWKFNVIIGHNDLCINNETTFGEETGKSLHS 215
 QY 537 ALDLHAEPRAFNALVTVEIIVMLRELYOEKV--YCPRMILRSICPVLLKPDNDSTEL 594
 Db 216 ALTIQTNVPKVFVN---IMPPIVAKHSQAHLKSKFC-EFSHKTKSCIFELNEKEYQ- 270
 QY 595 ATLIEFNKKOEKTHOLIE--SGRYDREDPTVVVQFFENVDMPKTSEGLPDNSFPAPD 652
 Db 271 ---NIKQPEQLNEVEQFNQKYGNSSTFAVIAIAMDLSIP-LIKQPNIGLALD 325
 QY 653 CEHSSKSHSAASALMNMMEPVGOKTRHKFNKINI-----TCPNQVQPIRTYKNS 707
 Db 326 CFHLSPYLAHDIAAKQIKWGLPEPIDQKT---ITNQSVCGRDRCVPCVBCFYLTIONS 381
 QY 708 MQGHGTWLPGRD---RAPSALHFTSVYALRPADIQV 740
 Db 382 ENCE----PSREMFLLVPSAFASPSGLSMPPWIPV 414

RESULT 7

; Sequence 6523, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 6523
 ; LENGTH: 348
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 ; US-10-369-493-6523

Query Match 5.2%; Score 406.5; DB 12; Length 348;

Best Local Similarity 32.2%; Pred. No. 3.2e-28; Indels 47; Gaps 11;
 Matches 117; Conservative 66; Mismatches 133;

QY 1060 IKPAIENWGSDFLCTEMKASNSV-----PTSVHQLRPADIKVVALGDSLTTAVGARNP 1113
 Db 15 VKQALE-----KILNSIDPHETIPDDVNMATKPHIRIYIGAMGDSL--IGSRAE 61
 QY 1114 NSSDLPTSMKGLSWSIGDGNLETHTLTPNLIK---KENPYLLGFSTSTWEG-TAGLVN 1168
 Db 62 NIVGQRORYPGNAFFPGMDPEVDRLTVVNIIFRIIAEKGKNGKLFGSTGIDYGENGLNV 121
 QY 1169 AAGCARAPMPAQMVLVERKSPDINLEKMKVTLTFGVND---LCHYCNPEAHILA 1225
 Db 122 AIGMKSDIDLRAKALVSRKANKKEINLEMDKVLSTWIGNDVGTGLGRDPIIP--V 179
 QY 1226 TEVVOHQIQAOLDLISEELPRAFNVVEVMEIASLYOGGCKAMLAQNNCTCLRSQSS 1285
 Db 180 DEYKSHIEKLLYKENTLPTIYSIYGMFAQLQDAQ-----SLKNGKRA 226
 QY 1286 LEKQELKXKNNMLQHGISEFSYHQYTQ---EDPAVVQPF-FQNTLPLNERGDTLT 1341
 Db 227 RIVENQKLD-DLSDGVRNVSYPQNNHENSNDFTVVQVQPFATETDYSRDEHGVNPT 285
 QY 1342 FPEEDCFHSDRCHAEKALYALMNMMEPVGKRTTSNNFHSRAKLKCPSPESPYLTLRN 1401
 Db 286 FVASDLFHLKSGFAVLAQKHYWNLPEPVGEKTRADLDGTKPKIYELNEKNCLIKTVGN 345
 QY 1402 SRL 1404
 Db 346 SKM 348

RESULT 8

; Sequence 6522, Application US/10369493
 ; Publication No. US20030233675A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Slater, Steven C.
 ; APPLICANT: Goldman, Barry S.
 ; APPLICANT: Chen, Xianfeng
 ; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 ; FILE REFERENCE: 38-10(52052)B
 ; CURRENT APPLICATION NUMBER: US/10/369,493
 ; CURRENT FILING DATE: 2003-02-28
 ; PRIOR APPLICATION NUMBER: US 60/360,039
 ; PRIOR FILING DATE: 2002-02-21
 ; NUMBER OF SEQ ID NOS: 47374
 ; SEQ ID NO 6522
 ; LENGTH: 349
 ; TYPE: PRT
 ; ORGANISM: Caenorhabditis elegans
 ; US-10-369-493-6522

Query Match

4.5%; Score 347; DB 12; Length 349;


```

;      :      :      :      :      :      :      :      :      :      :
Db      NSMALQKITLQFAHFLIEIDSPSLKTEILEDFLLVTKMLOEYANEDYSRMETLPI-- 646
Qy      589 CULTRINSEGLARLEAFSAAYRSMRELYVSGCYD-----TOE---DFSVYLOPF 978
;      :      :      :      :      :      :      :      :      :      :
Db      647 -PVTNESSTEDIALALAVAIATFWGSLKNISREGNFDVAFLTHLNOQLTNFSVV-QLLF 704
;      :      :      :      :      :      :      :      :      :      :
Qy      979 QNTQLPVL-----ADGLPPTSEFPAPDCIHPNQKFSQSLARALMTNMLEPLGSKTE 1028
;      :      :      :      :      :      :      :      :      :      :
Db      705 ENLILNINNLNLAGNSQEAANNLNTDQIMNFN-----LILNHMQSETSRKT 752
;      :      :      :      :      :      :      :      :      :      :
Qy      1029 TLDLRAEMPTTCPTONEPFLRT 1050
;      :      :      :      :      :      :      :      :      :      :
Db      753 VLSRSIVDFT-----EQFLKT 769

```

```

RESULT 11
US-10-156-761-13312
; Sequence 13312, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMIURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156, 761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13312
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-13312

```

```

Query Match      1.8%; Score 136; DB 15; Length 300;
Best Local Similarity 27.3%; Pred. No. 0.0022;
Matches 51; Conservative 31; Mismatches 71; Indels 34; Gaps 9;

```

```

Qy      1083 PTVSHQLRPADIKVVAALGSLTTAVGARPNNSSDLPTSWKGLSWSIGDGNLETHHTLP 1142
;      :      :      :      :      :      :      :      :      :      :
Db      48 PTVWMDRSPASL---AAVGDSITRGPDACA-VLSDCE---VSMATGSDADVDLSLA--- 96
;      :      :      :      :      :      :      :      :      :      :
Qy      1143 NILKKFNPLYLGSTSTWEGTAGLVNAABGARADMPAQAMDLYERKMSPDINLEKDWK 1202
;      :      :      :      :      :      :      :      :      :      :
Db      97 --VRLGSP--VGAERSW-----NYAATGARMAADLPEQMERAVARRP-----E 135
;      :      :      :      :      :      :      :      :      :      :
Qy      1203 LVTLFTGVNDLCHYCEPBEAHLATEVYQHIOQLDISELPFAFVNVVEVMEIASLYOG 1262
;      :      :      :      :      :      :      :      :      :      :
Db      136 LVTVWGANADACRASASMTSVA--DFRADLEDAVATVQALPKIQVTMSVVPNLKRLM-S 193
;      :      :      :      :      :      :      :      :      :      :
Qy      1263 QGGRKAM 1269
;      :      :      :      :      :      :      :      :      :      :
Db      194 QGRTSAL 200

```

```

RESULT 12
US-10-369-493-1156
; Sequence 1156, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.

```

```

; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369, 493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360, 039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 1156
; LENGTH: 1708
; TYPE: PRT
; ORGANISM: Methanobacterium thermoautotrophicum
US-10-369-493-1156

```

```

Query Match      1.7%; Score 133; DB 12; Length 1708;
Best Local Similarity 20.7%; Pred. No. 0.1;
Matches 127; Conservative 83; Mismatches 202; Indels 202; Gaps 29;

```

```

Qy      593 ELATLIEFNKKFQDKT--HQLIESGRYDREDFTVVVQPFENVMPKTSGLPDNSFFA 650
;      :      :      :      :      :      :      :      :      :      :
Db      181 DLTSILAYKOKYKGNALHOMIDICALYVAAAGKTINLEOPKALKMYNNRGILPMSSWE 240
;      :      :      :      :      :      :      :      :      :      :
Qy      651 PDCFHSSKSHSRAASALMNMNMLEPVGQKTRHKFKFNKINITCPNOVQFLR----- 702
;      :      :      :      :      :      :      :      :      :      :
Db      241 P-----ATVPASPLSEFLYRDQRTKEDYFTRVPLDPAKPTVAVLVYSGTGEV 292
;      :      :      :      :      :      :      :      :      :      :
Qy      703 TYKSMQHGHTWLPCHDRAPSAHPHPSVHALRPADIQVVAALGDSLTAAGIGISKDDLP 762
;      :      :      :      :      :      :      :      :      :      :
Db      293 TYADAMQ-----QIDALVSRGLNVLPIVIG--TWSYV----- 353
;      :      :      :      :      :      :      :      :      :      :
Qy      763 DVTTQYRGLSYSGDGSLENVTLLPNILREF-----NRRLGYAVGTGDANDTNALN- 816
;      :      :      :      :      :      :      :      :      :      :
Db      324 -----ILNQGMQNIIOQLCMPQNTNITAIR-GIGNYDPLTSLGV 364
;      :      :      :      :      :      :      :      :      :      :
Qy      817 QAVPGAKABDLMSQVQTLMOKKMDHRVNFHEDEKVIYVLIGSGLCDVCTDSNLYSAAN 876
;      :      :      :      :      :      :      :      :      :      :
Db      365 TSVSAN--VYEQIL-----DNGNLIIRSLKISTV-----QPVNVYSA-- 400
;      :      :      :      :      :      :      :      :      :      :
Qy      877 FVDHARNALDYLHREVPRLVNLVDFLNPIMQVFLGNPDKPVQOASVLCNCVLTRE 936
;      :      :      :      :      :      :      :      :      :      :
Db      401 LVKFLTDASNVQYE-----ANPEKYPV-KANVITDMLTFTTG 437
;      :      :      :      :      :      :      :      :      :      :
Qy      937 NSQELARL-EAFRAVRSSRMRELVGSGRYDT-----OEDFSV-----LQPFQNIQ 982
;      :      :      :      :      :      :      :      :      :      :
Db      438 STTSGASVTKFPRKSNIPVLKAMITSTYRTTIGQTVSESGFMSVYMQCAPEHQG-Q 496
;      :      :      :      :      :      :      :      :      :      :
Qy      983 LPTVADGLPDTSEFPAPDCIHPNQKFSQSLARALMTNMLEPLGSKTETLDRAMPITCPT 1042
;      :      :      :      :      :      :      :      :      :      :
Db      497 IETLAIQVGE-----IGSDPET-GAQMWDITVTYPE 525
;      :      :      :      :      :      :      :      :      :      :
Qy      1043 QNEPF-----LTPRNSN-----YTYPIKAIEWGSDFLCTEKANSVPTSV 1086
;      :      :      :      :      :      :      :      :      :      :
Db      526 RIEKLVSRAFNMIIRLQTMANSDDKVAIVYVNP--PGKQNIIGSYL-----NVPEST 575
;      :      :      :      :      :      :      :      :      :      :
Qy      1087 HOLRPADIKVVAALGSLTTAVGARPNNSSDLPTSW--RGLSWSIGDGNLETHHTLPN- 1143
;      :      :      :      :      :      :      :      :      :      :
Db      576 IET-----LKKMAEGIS-----VGEIQDADALVEMIKGIVANNAAPGELEKLANSSNA 627
;      :      :      :      :      :      :      :      :      :      :
Qy      1144 ILKKFNPLYLGFT 1157
;      :      :      :      :      :      :      :      :      :      :
Db      628 ILWPYEDYLAWFNT 641

```

```

RESULT 13
US-10-359-012-4
; Sequence 4, Application US/10359012
; Publication No. US20030232419A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: KOLODKIN, Alex L.
; APPLICANT: TERMAN, Jon R.
; APPLICANT: MAO, Tianyi

```

APPLICANT: PASTERKAMP, Ronald J.
 APPLICANT: YU, Hung-Hsiang
 TITLE OF INVENTION: MOLECULES INTERACTING WITH CASL (MICAL) POLYNUCLEOTIDES, POLYPEPT
 TITLE OF INVENTION: AND METHODS OF USING THE SAME
 FILE REFERENCE: JHU1840-3
 CURRENT APPLICATION NUMBER: US/10/359,012
 CURRENT FILING DATE: 2003-02-04
 PRIOR APPLICATION NUMBER: US 60/388,325
 PRIOR FILING DATE: 2002-06-13
 PRIOR APPLICATION NUMBER: US 60/384,302
 PRIOR FILING DATE: 2002-05-30
 PRIOR APPLICATION NUMBER: US 60/354,178
 PRIOR FILING DATE: 2002-02-04
 NUMBER OF SEQ ID NOS: 40
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 4
 LENGTH: 1633
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-359-012-4

Query Match 1.7%; Score 129.5; DB 12; Length 1633;

Best Local Similarity 19.3%; Pred. No. 0.2; Matches 222; Conservative 118; Mismatches 424; Indels 385; Gaps 49;

18 OGTPQHTSPRKSTLEGOIWPETLKNSPPPCNPNK-----LGVMPSKSHSLKPSD 69
 430 OGTPPELAEERSLV-RILPOTT-----PENINKPEOYTLDPGTRYPYLNHCHVAFHQ 483
 70 IKFVAIAGNLEIPDPDGTGLEK-----QDWTERPOQCMGVT- 108
 484 VKLVTYTKLEHYPLERTGSVRSVNLRSKESDIRPSKLLTWQOQTEGQVHNVDLTT 543
 109 -----VLSIIIRYFSPSV-----PMPVCHTKGVYIEMD 136
 544 SWESGALCALIHRFRPELINPDSLNEDEAVENNOALFVDAEREFGI.PPVTTGKEMASQA 603
 137 GAEDLMT-----OAOELVR-----NMKENQLDPQFDWKLINVFASQCYL 179
 604 EPKLSMVMWLSKFYELFRCTPIRAPVDSMKKNYGENADLS-----L 644
 180 CPBAQNGLAAGVDELMGVLDYLOQEVPRAPVNLVD-----LSEVA 221
 645 AKKSISN-----NYLNLTPRKTRPVDCQGTGENDMNRKRKFTNIDEPS 690
 222 EVSRQYHGTWLSAPBPCNSEETTRIAKVWQMSVOEANNLSIAS-----SRVSEQ 273
 691 NFSSRSISG-----SNOECSSKEGQNKV-----KSMANQLLAKFEESTRNPSLMKOE 739
 274 ESTTVFQPPFFYETT-PSLHSEDRLODST-----TLAMHLMNMEPAGE 318
 740 RRVSGIGKPVLCSSSGPVHSCCPKEAPSPSPPLKQFPSSVVYGAHL-BELLQVSA 798
 319 KOEPLSVKIGRPMKCSQESPVLFYSYN-----SNVYLRLOKPODKLEVR 364
 799 GSECLG-----RPWRARAKSDLOLQGTENFATLSTRPRAALSGVLMRLQOVEKILQKR 854
 365 GAEIRCPDSDPTVPTSVHRLKPADINVIAGLSDSLTAGNGAGSPFGNVLVDLTOYRGL 424
 855 AOMLANREFTHKIKEXAAH-----LASWFGHGDPPQNTL-----LSKGL 894
 425 SMS-----VGDENIGTYTTIANILREFPISLKGFSVGG-KETSPNAPLN 469
 895 SHHPPSPSPRLPDPDPAASSSPSTVDS-ASPAKKEKSSPGHFHSHLRTVHP----- 948
 470 QAVAGGAEADLPYQARLYVDMKNDRIFHQED-----WKITITLFIGGNDLDFCNDL 522
 949 QLVGVKSSGIGAAAEVLVNLVNDHRPKQATSPDLESWRKSPPLNLGSSDTCYCKKR 1008
 523 VHSFQNTDNIGKALDILHAEPRAFVNLVTVLEIVNRELVOEKKVYCPRMILSLCP 562
 1009 VY-----VMERLTAEGHFFHRECFRCST-CATTLRLAAYTFDCEGKFYCK-----P 1054

583 CVLKPDNDSTELATLIEFNKKQOEK-THQILBSGRYDREDFTVVVQPPFENVDMPTSE 641
 1055 HFHICTNKKQRRARALKQOREEATWQOEAPRDTPTSECAVAAL-----CTLE 1107
 642 GLPDNSFPAPDCRHSSKSHRAASAL-WNNMLE-----PVGCKTRHKFNKINTTCPN 695
 1108 GSPPVHFSLP-----VLHPLGLMLDWNDSIPESVHLKAGERISQKSAENGRGRVLK 1160
 696 QVQPFRTYKNSMGHGTWLPCCDRAPASLHPFSVHALRPADIQVVAALGDS----- 747
 1161 PVABLL-----LP--RAGEPLPTQRAQOEKMGTRPAOQOERVPPPKSP 1204
 748 -LTAGNGISKPPDLPDVTYQYKGLSYSAGD--GSLVENTTLPN-----TLRENNR 797
 1205 LRLIANAIRSLLEPL-----LNSBGKKAMAKQESKITPLQACTRSFSLRKTNSN 1255
 798 LTGYAVGTGDANDTNAFLNQAAPGAKAEDMSQVOTLMQMKDHRVNFHEMKVITVLI 857
 1256 KDG-----DQHSFGRNQSSAFSPDPALRTHSLRNPSS-----KVFPAL- 1294
 858 GSGDLCDCYCTDSVLYSAANFVDHLRNALDVLHREVRVLVNLVDPLNPTMRQVFLGN-- 915
 1295 -----RSPCKSIEEV-----PTLKVSLQENF 1318
 916 PD--KCPVOOASVLCNVCVTLRENSQELARLEAFSAHYSMSRELVSGR-----YDQE 968
 1319 PDASKPKKRIKSLFSS--LRKDKSFESPLQES--RQKRDITDLGSPRKVLPEDSAQ 1373
 969 DRSVLQPF 977
 1374 ALEKLLQPF 1382

RESULT 14

US-09-808-602-84

Sequence 84, Application US/09808602

Patent No. US20020155115A1

GENERAL INFORMATION:

APPLICANT: Vernet, Corine A

APPLICANT: Fernandez, Elma

APPLICANT: Shimkels, Richard A

APPLICANT: Herman, John L

APPLICANT: Majumder, Kumud

APPLICANT: Mishra, Vihenu

APPLICANT: Mezes, Peter S

APPLICANT: MacDougall, John

TITLE OF INVENTION: No. US20020155115A1el Proteins and Nucleic Acid Encoding Same

FILE REFERENCE: 15966-697 CIP

CURRENT APPLICATION NUMBER: US/09/808,602

CURRENT FILING DATE: 2001-03-14

PRIOR APPLICATION NUMBER: 09/800,198

PRIOR FILING DATE: 2001-03-05

PRIOR APPLICATION NUMBER: 60/186,596

PRIOR FILING DATE: 2000-03-03

NUMBER OF SEQ ID NOS: 114

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 84

LENGTH: 2765

TYPE: PRT

ORGANISM: Rattus norvegicus

US-09-808-602-84

Query Match 1.7%; Score 128.5; DB 10; Length 2765;

Best Local Similarity 17.9%; Pred. No. 0.66;

Matches 277; Conservative 194; Mismatches 511; Indels 565; Gaps 71;

256 SYOEANSLASRSYSESPFVVPQPPYETTPSHSDPR-----LQDSTLAWH 307
 1112 AITFTWKTDAYGQRYGJSDAVSVGFYETCPSLILWEKTKALLQGFELDPNSIGWS 1171
 308 LWNRMMEPAGEKDEPLSVGHG-----RPMKCPG----- 335
 1172 L-----DKHHTLVNVSGLLKGTEGNOPLTQOPAIITISMGNGRRRSISCPCNGL 1222

```

QY 336 QESPYLFS-----YRNSNYLTRLOKPOD-----KLEVEGAEIRCPDK----- 373
DB 1223 AEGNKLLAPVALAVGIDGSLFVGDFNYIRIRIPSRVNTSILELRNKEFKHNSPGHKYYL 1282
QY 374 -DP-----SPTVTSVHRLKPADINVIAGLDSLTAGNAGSTEGNVLDTLYOYRGL 424
DB 1283 AVDPVTGSLYVSDTNSRRIRYRK-----SL-----SGAKDLAGNSEVV----- 1320
QY 425 SMSVGDENIGYTTTLANIILREFNPBLKGFSGVGTGKTSBPNAFLNQAVA--GGRADLPV 482
DB 1321 -----AGTGEQCLP--FDEARCGDGGKAVADATL 1346
QY 483 QABRLVDMKND-----TRHFOEDMKITLFTIGNDL-----CDFCNLDVHSPQ 528
DB 1347 MSRGIAVDCXGLMYFVADMIKRVNDNGIISTLGSNDLTAARPLSCDSMDVAQVRLE 1406
QY 529 NFTDNIGKALDILHAEPVRAFNVLVYLE--IVNLRELYQ--EKVYCC----- 572
DB 1407 WPTDLAVNPM-----NSLYLENVILIRITENHOVSIIAGRPMHCQVPGIDYS 1455
QY 573 -PMLRLSLCPCLKPDNSTELATLIEFNKKFOEKTHQILBSGRY-----DTRB 621
DB 1456 LSKLAHSALESASAJAISHTGVLVITETDEKKNRLROYTTNGEICLLGAASDCCKN 1515
QY 622 DFTVYVQPFEN-----VDMPKTSEGLPDNSFPAPDCFPSSKSHR----- 663
DB 1516 DVNVCICSGDAATADAIINSFSLAVAPDGTIYIADLGNIRRAVSKKAPVLAFFOYE 1575
QY 664 AAS--ALMNNMLEPVQOKT--TRHKFNKINITCPNOVOPPLRTYKNSMOGHGMLP 716
DB 1576 AASPEGELEYVFNADGHQYTVSLVTEGVLYNFYTSADNDVTELDINNGSL----- 1627
QY 717 CRDAPBALHPTSVHARPADIQVVALGSLTRAGNGISGKPDLPVYTTQ-----Y 768
DB 1628 -KLRDSSGNMR--HLLMP--DNOI--TLVTGTNGGKA-----VSTONELGLMTY 1673
QY 769 RGLS-----YSAGDGSLENT--TLPNILREFNNLGVAVGTGDAN 809
DB 1674 DGNITGLATSDDETGWTFYDYHREGRLTVTRPTGVVTSILHEMESISVIDIENSNRDN 1733
QY 810 DTNAFLNQAVGAKAEDLMSQVQ-----TLMQKMDHRYNFHE----- 848
DB 1734 DVAVITLNLSEVSEAYTVVQOVNRNSYQLSNGTLRVAVYANGMGVSPSEPHVLAGTLPT 1793
QY 849 -----DMKV-----ITVL-----IGSDLCYCTSNLSAANFV 878
DB 1794 IGRCNISLPENGLINSIEWRLKEQIKGKVTIFGRKLRVHGRNLSIDYDRNIRTEKIYD 1853
QY 879 DHLRNALDVHREYPR-----VLNVLVDFLNPFTIKRQVFLGNPDKCPVQOQSVLC 928
DB 1854 DHKFTLRITTYDOVGREFLMLPSSGLAAVAVSYFFENRGLGLQRGAMSEETDIDKOG-- 1910
QY 929 NCVLTLRENSOEALRAEFSPRAY--RSSMRELVSGR-----YTOEDFSVVLQ-- 976
DB 1911 -----RIVGRMPFADGVMSYSYLDKSMVLLDSQORQYIFREYSSDRLLHVTMPVARHS 1964
QY 977 -----FQONQLP-----VLADGLPD--TSFFAPDCIHPQKH-----SOLA 1012
DB 1965 MSTHTSIGYIRNIYNPESNASAVIFDYSDGRILKTSFLGTG--RQVYFKGKLSKLS 2020
QY 1013 BALM-----TNMLEPLGSKTETLDLRAEMFITCPQNEPFLRPRNS----- 1054
DB 2021 EIVYDSTAVTFGDETTGLMVMNLQSGFSCITIRYKVGVLVDKQIYRSEBGMINARF 2080
QY 1055 NYTY-----PIKPAI-----ENMGSDPLCTEMKASNSVPTSVH 1087
DB 2081 DYTYHDNSFRIASIKPIVSETPLPVLDLYRDEISGKKEHFK--FGVLYYDINOITTAVM 2139
QY 1088 QL-----RPADIKY-----VALAGDSLTTAV-----GARPNNSDLPISWRG 1124
DB 2140 TLKSHPDTHGRIKVEYEMFRSLMYMTVOYDSWGRVYIKRELKLGPAANTTK----- 2191

```

```

QY 1125 LSWSIGDGNLETHHTLTPNLLKKNFNYLGFSTSTWEGTAGLN-----VAEGARARDMPA 1180
DB 2192 YTYDVYDGDGLQ-----SVAVNDP-----TWRYSDLNGNHLNPGNSARLMLP 2237
QY 1181 QANDLYERKNSPDINLEKD-----WKLVTLPFGY 1210
DB 2238 R-YDLRDRITRLDVOYKIDDDGYLCQSGSDIFEVNSKGLLTRAANKASGWSVQYRYDGV 2296
QY 1211 NDLCHYCENPEALHNEY--VOHIQOALDILSELPRAFVNV--EVMELASLY----- 1260
DB 2297 SRPASTYKLNGLHLLQYFSDLH-----PRTITHYVNSNETTSLYYDLOCH 2344
QY 1261 ----OGGCKAMLAQNNCTCLRHSSQSLKEOKELKRVNMLQHGISFSPYHQYQRED 1316
DB 2345 LFAMESSGSEYVVASDNGTLP--AVFSINGLMIKQLOYTAVGEI-----YDSONPD 2395
QY 1317 FAVVOPFFONTLTPLNEBGDDLTFTFSBDCFFSFGHAENALALMNNMLEPVGKRTTS 1376
DB 2396 FQWVI--GFHGLYDPLTK-----LVHFTORDVDVLA--GR-WTS 2430
QY 1377 NNTSHRAKLKCPSPESPLYTLRNSRLPDQAE--EAPVLYMAV 1420
DB 2431 PDYTMWRNVGKEPAPFN--LYMEKNNPNLSNELDLKNVYTVDKSWLV 2475

RESULT 15
US-09-800-198-72
; Sequence 72, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Verne, Cornie AM
; APPLICANT: Fernandes, Blma
; APPLICANT: Shinkete, Richard A
; APPLICANT: Hejmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 72
; LENGTH: 2765
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-800-198-72

Query Match 1.7%; Score 128.5; DB 11; Length 2765;
Best Local Similarity 17.9%; Pred. No. 0.66;
Matches 277; Conservative 194; Mismatches 511; Indels 565; Gaps 71;

QY 256 SYEANSLSLASRYSEOSTFYVFPFFYETTPSHSDPR-----LQDSTLAWH 307
DB 1112 AYTFIDKTDAGQRYGSLDAVAVGVFEYETCPSLILWEKRTALLQGFELPNSLGWS 1171
QY 308 LMRMMEPAGEKDEPLSVKHG-----RPMKCPG----- 335
DB 1172 L-----DKHITLVASGILLKGTGENQPLTOOPALITTSIMGRRRSISCPSCNGL 1222
QY 336 QESPYLFS-----YRNSNYLTRLOKPOD-----KLEVEGAEIRCPDK----- 373
DB 1223 AEGNKLLAPVALAVGIDGSLFVGDFNYIRIRIPSRVNTSILELRNKEFKHNSPGHKYYL 1282
QY 374 -DP-----SPTVTSVHRLKPADINVIAGLDSLTAGNAGSTEGNVLDTLYOYRGL 424
DB 1283 AVDPVTGSLYVSDTNSRRIRYRK-----SL-----SGAKDLAGNSEVV----- 1320
QY 425 SMSVGDENIGYTTTLANIILREFNPBLKGFSGVGTGKTSBPNAFLNQAVA--GGRADLPV 482

```

Db 1321 -----AGTGECLP-----FDEARCGGKAVADATL 1346
Qy 483 QABRLVDMKND-----TRHFQEDWKITLFIGNDL-----CDPCNDLVHNSPQ 528
Db 1347 MSBREGAVDNGMLYFVADATMIRKVDONGIITLLSNDLTAVRPLSCDSMVAOVRLE 1406
Qy 529 NFDNIGKALDILHAEPRAVNLVTVLE---IVNIRELYO-----EKRYC----- 572
Db 1407 WPTDLAVNPD-----NSLYLENNVILRITENHOVSIIAGRPNCQVPGDYS 1455
Qy 573 -PRMILSLCEPVLCRDKDNDSTELATLIEFNKFOEKTQOLIESGRY-----DTRE 621
Db 1456 LSKLAHSALESASALAIASHGTVLYTETDEKINRLRQVTNGEICLGAASDCCKX 1515
Qy 622 DFTVVVQPFEN-----VDMPTSEGLPNSFFAPDCFFSSKSHR----- 663
Db 1516 DVNCCYSGDADAVATDALINSFSLAVAPDGTIYIADLGNIRIRAVSKNRPVLANAFNOYE 1575
Qy 664 AAS-----ALMNNMLBPVGOKT---TRHKFNKINITCPOVOPLRTYKSMOGHTWLP 716
Db 1576 AASPGOELYVFPADGHOYTSLVGEYLYNFTYSADNDVTELIDNNGSL----- 1627
Qy 717 CRDRAPSAHPTSVHALRPADIQVVAALDSLTAGNGISGKPDLPDVTQ-----Y 768
Db 1628 -KIRDSGMPR--HILMP-DNQII-----TLTVGTNGGLKA-----VSTQNLGLMTY 1673
Qy 769 RGJS-----YSAGDGSLENT-----TLPHILRFPNRLTGIVAGTGDAN 809
Db 1674 DGNITGLATKSDETGWTTFDYDHEGLTNTVTRPTGVVTSLHREMEKSIIVDIENSNDN 1733
Qy 810 DTAFLNOAVPAGAKEDLSMOVO-----TLMOCKMDHVRNPFH----- 848
Db 1734 DTVVITNLSVEASYVVOQVNSYOLCNGTLRWYANGMGSFHSSEPHVLAGTILPT 1793
Qy 849 -----DMKV-----ITVL-----IGSDLCYCTDSNLYSANFV 878
Db 1794 IGRCNISLPENGLNSIEWRLRKEQIKGKVTIFGRKLRVHGRNLSIDYDRNIRTEKIYD 1853
Qy 879 DHRNALDVLRHVR-----VLVNLVPLNFTIRQVFLGNDPKCPVQASVLC 928
Db 1854 DHRKFTLRITIDOVGRPFLWPSSGLAAVNVSYFFNGRLAGLORGAMSERTDIDKOG-- 1910
Qy 929 NCVLTLRENSOELARLEAFSRAV--RSMRELVSGR-----YDQEDFSVLOP----- 976
Db 1911 -----RIVSRMPADGKWSYSTLDKSMVLLQSOROYIREFYDSSDRLAHVMPVSARHS 1964
Qy 977 -----FQONIQLP-----VLADGLPD-----TSFFAPDCIHPNQKH-----SOLA 1012
Db 1965 MSHHTSIGYIRNYNPPESNASVIFDYSDGRILKTSFLCTG---RQVYKXGKLSKLS 2020
Qy 1013 RALW-----TNMLEPLGSKTETLDLRAEMPTCPTQNEPLRTPRNS----- 1054
Db 2021 EIVYDSTAVTFGYDETTGVLMVNLQSGFSCIRYRKVGPLVKQIYRFESEGMINARF 2080
Qy 1055 NNTY-----PIKPAI-----EMWGSDELCTEMKASNSVPSVH 1087
Db 2081 DTYHDNSFRIASIKPIVISETLPVLDLYRYDEISGKVEHFGK--FGVLYYDINOIITTAVM 2139
Qy 1088 QL-----RPADIKV-----VAALGSLTTAV-----GARPNNSDLPTSMRG 1124
Db 2140 TLSKHFDTHRIKEVOYEMFRSLMYMTVOYDSMGRIKRELGLGPIANTTK----- 2191
Qy 1125 LSHSIGDGNLEHTHTLPLNLKKFNPYLLGFSTSTWEGTAGLN---VAAEGARADMPA 1180
Db 2192 YTYDYDGDGLQ-----SVAVNDRP-----TWRYSYDLNGNLHLNPGNSARLMP 2237
Qy 1181 QAMDLYERMKNSPDINLEKD-----WKLVTLPFGY 1210
Db 2238 R-YDLRDRITRLDVOYKIDDDGYLCQSGSDIFEYNSKGLLTRAYNKASGWSVOYRYDGV 2296
Qy 1211 NDLCHCENPEAHATEY--VOHIQCALDILSEELPRAFNVV--EVMELASLY----- 1260

Db 2297 SRRASYKTNIGHNLQYFYSDLHH-----PTRITHVYHNSSEITSLYYDLOGH 2344
Qy 1261 ----OGGKCAMLAQNNCTCLRHSQSLKEQELKXVMNLIQHGISFSYMHQYQREP 1316
Db 2345 LFAMESSGEBEYVVASDNTGTPU--AVFSINGMLTKQLOTAYAGEI-----YDSNPD 2395
Qy 1317 FAVVVOFPFONTLPLNERGDTDLTFPSDCPHFSDRGHAEMALMNNMLBPVGRKTS 1376
Db 2396 PQWVI-GFHGLYDPLTK-----LVHFTQRDYDLA-----GR-WTS 2430
Qy 1377 NNFTSRAYLKCBSPSPYLYTLRNSRLLPDOAE---EAPVLYWAV 1420
Db 2431 PDYTMWRNVGKEBAPFN--LYMFKNNNPLSNBLDKNYVTVDVKSMLV 2475

Search completed: January 6, 2004, 19:05:43
Job time : 379 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 6, 2004, 18:59:25 ; Search time 31 Seconds
(without alignments)
1989.975 Million cell updates/sec

Title: US-10-054-691-2

Perfect score: 7766
Sequence: 1 MGRPCIFLEILLIGQGT.....RCRRGRREDPPMSRLTVAL 1458

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 4231058 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patente_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	138	1.8	10182	4 US-09-134-001C-3159	Sequence 3159, App
2	136	1.8	1455	2 US-08-726-012B-2	Sequence 2, App1
3	129.5	1.7	1218	4 US-09-589-567-2	Sequence 2, App1
4	127.5	1.6	1016	3 US-09-180-439-8	Sequence 8, App1
5	123.5	1.6	2636	4 US-09-252-991A-25753	Sequence 25753, App
6	122	1.6	855	4 US-09-107-532A-5646	Sequence 5646, App
7	121.5	1.6	1112	3 US-09-353-585-2	Sequence 2, App1
8	120.5	1.6	2842	1 US-07-741-940-7	Sequence 7, App1
9	120.5	1.6	2842	1 US-08-289-548A-7	Sequence 7, App1
10	120.5	1.6	2842	1 US-08-452-654-7	Sequence 7, App1
11	120.5	1.6	2842	4 US-08-449-731-7	Sequence 7, App1
12	120.5	1.6	2843	1 US-07-741-940-2	Sequence 2, App1
13	120.5	1.6	2843	1 US-08-289-548A-2	Sequence 2, App1
14	120.5	1.6	2843	1 US-08-452-654-2	Sequence 2, App1
15	120.5	1.6	2843	1 US-08-452-655B-2	Sequence 2, App1
16	120.5	1.6	2843	1 US-08-452-655B-7	Sequence 2, App1
17	120.5	1.6	2843	2 US-08-370-235A-2	Sequence 7, App1
18	120.5	1.6	2843	3 US-08-450-582-2	Sequence 2, App1
19	120.5	1.6	2843	3 US-08-450-582-7	Sequence 2, App1
20	120.5	1.6	2843	4 US-08-449-731-2	Sequence 7, App1
21	120.5	1.6	2973	2 US-08-821-355A-7	Sequence 7, App1
22	120.5	1.6	2973	2 US-09-003-687A-7	Sequence 7, App1
23	120.5	1.6	2973	3 US-09-136-605-7	Sequence 7, App1
24	120	1.5	15281	2 US-08-471-119A-2	Sequence 2, App1
25	119.5	1.5	1112	3 US-09-353-585-3	Sequence 3, App1
26	118	1.5	1007	3 US-08-961-083-216	Sequence 216, App
27	118	1.5	1007	4 US-09-536-784-216	Sequence 216, App

28	117.5	1.5	956	4 US-09-134-078-63	Sequence 63, App1
29	117.5	1.5	2509	4 US-09-252-991A-16642	Sequence 16642, A
30	116	1.5	1430	2 US-08-540-804-14	Sequence 14, App1
31	116	1.5	1430	2 US-08-218-265-14	Sequence 14, App1
32	116	1.5	1430	3 US-08-521-872-14	Sequence 14, App1
33	116	1.5	1430	3 US-08-590-399-14	Sequence 14, App1
34	116	1.5	1589	3 US-09-356-952-4	Sequence 4, App1
35	114.5	1.5	1447	4 US-09-376-330-17	Sequence 17, App1
36	114.5	1.5	2154	2 US-08-841-349-4	Sequence 4, App1
37	113.5	1.5	968	3 US-09-180-439-4	Sequence 4, App1
38	113.5	1.5	968	3 US-09-180-439-4	Sequence 4, App1
39	112.5	1.4	1180	3 US-09-224-024-28	Sequence 28, App1
40	112.5	1.4	1180	5 PCT-US94-07902-28	Sequence 28, App1
41	112	1.4	1018	5 PCT-US93-11703-75	Sequence 75, App1
42	111.5	1.4	797	3 US-09-086-912-2	Sequence 2, App1
43	111.5	1.4	797	4 US-09-203-453-2	Sequence 2, App1
44	111.5	1.4	797	4 US-09-500-236-2	Sequence 2, App1
45	111.5	1.4	951	3 US-08-816-346-58	Sequence 58, App1

ALIGNMENTS

RESULT 1
US-09-134-001C-3159
; Sequence 3159, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134, 001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064, 964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055, 779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3159
; LENGTH: 10182
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3159

Query Match 1.8%; Score 138; DB 4; Length 10182;
Best Local Similarity 18.4%; Pred. No. 0.1;
Matches 265; Conservative 187; Mismatches 546; Indels 440; Gaps 61;

QY 153 KENLDQDFQEDWTLINVFSSNASQCYLCPSAONGLAAGVDEL-----MGVLDV 202
DB 4346 KNNLDVDTPLKXIKKQLDEIDQ-----GTNTDGTQDSDVDVNDLSAIIKKKVK 4400
QY 203 LQEEVRAFNLDLSEVAEVSROYHTWLSAPAPCNSSEETTRAKVVMQSYOBA-- 260
DB 4401 LKRNPTGVQVESVANAQOVIQDLOMARTSLVDPDQLOEAGNRLENSIQOTDIDGKT 4460
QY 261 -----WNSILASRYS-BOESFTVFQPFYETTPSLSEBRLDSTTLAVHL-WNRM 312
DB 4461 QDSLNMYNKKLAKARONLEKISKVLGGQPTVAIRQNTDEANNAHQALDTPARQLTLNR- 4519
QY 313 MEPAKEDPLSVKXHKRPMKCSQSPYLPSTYNSNYVLRQKPRQKLEVBREAEIRCPD 372
DB 4520 -----EPYINHINNESHLNNAQDNFKAQVNSAPNNTLE 4554
QY 373 --KDPEDVPTSVHRLKP--ADI-----NVIGA-----LGDSLTAGNG--AGSTP 411
DB 4555 TIKNRKDTLNGSTALSESIALDYENKQOENYLDASNNKRODYDAVNAAKGILNQTSF 4614
QY 412 GNVLDVLTQ---YRGLSWSGDENIGVTYTLA-----NLRFNPSLKFSGVGTGKTS 463
DB 4615 TMSADVIDQKAEDVKRTKTRTALDQNGRLBVAQOALNHLNTLNDLNDQAGQTLTDITNH-S 4673

```

QY 464 PNA-FLNQAAGRAEDLPVQARLVLMKN-DTRHFOEDMKTITLFIGNDLDCFCND 521
DB 4674 PNINSVQAQAE--KAANTVNTAMTOLKOTIANYDELH---DGNITN---ADKXKDXANN 4725
QY 522 LVHYSFONFTDNIKALDILHAEPRAVNLVLEIVNLRELYOEKKVYCCPRMILSLC 581
DB 4726 AVNNAKOLINQSDANOQALDPAEINKYQVRVNTTKNDLNGDKLAEAK----- 4773
QY 582 PCVLKPFNDNSTELATLIEFNKKFQEKTHQULESRYOTREPFYVVOGPFPE-NYDMKTS 640
DB 4774 -----RDANTTIDGLTYLLEAQNNKAKENV--GKASTKNTITSQLODYNOQNTAMQALR 4825
QY 641 EGIJPD-----NSFPAPDCFHFSKSHSRASALMNMMLP-----VGOKT--- 680
DB 4826 NSVVDVANNVANSYINEDNPKKANYQAVTH---AQTLINAGSNPMSRUVNOKQAV 4882
QY 681 -TRH-----KFE-----NKINTT-----CENOVOPFLRTYKN 706
DB 4883 NTAHQNLHGQOKLEQAQSSANTEIGNLPNLTNTQAKKEKELVNSKQRTVEQOLNQAKS 4942
QY 707 SMOGHQWLPORDBAPBALHTSYHALRPADIQVVAALGSLTGNGIGSK-PDDLDDVT 765
DB 4943 LDSSWGLKSLVAQOPT-VQKTSYVINEDQEPQ--SAVNSITMGQTTINKTADPVLDKT 4999
QY 766 TOYRGLSYSAGDGLSNVTTLPVILREPNRLTGVAVGTGDANDTAFNL---QAVPG 821
DB 5000 LV-----DNAISNISSTKENALHG-EOKLTAKTEANALNTLADLNTPOKKAIKT 5048
QY 822 A-----KAEDLMSQVOTLMQMKDHRVNFHEWKVITVLIGSGDLCDYC 866
DB 5049 AINTAHTRTDVAQSKANOINSAMHTLRQNISDNESVT----- 5087
QY 867 TDSGLYSNAA-----NFPDHLRNALDVLHREVRVNLVLPFLNTIRQVFLGNPKCPV 921
DB 5088 NESYVINAEBEKQHAFTALNNAEIVNEQATLIDANSIN----- 5127
QY 922 QQASVLCNCVLT-----RENS--QELARLEAFSRAVSSMRELVSGRYDTQ--- 967
DB 5128 QKQAOILITTKNALDGEOLBRAKENADOEINTLQOLTDQORNSKGLINSQTTTEVASQ 5187
QY 968 -----EDFSVYLQPFQONIQLEVLADG---LPDTSFPAEDCIHPNQKFSQALRALWMTML 1020
DB 5188 LAKAKEINKWME-----QLNHLINGKNQWINSKFINEDANQOQAVSNMIAAS----- 5235
QY 1021 EPLGSKETITDLRAEMPTCTQNEPFLRTPRNSNYTYPYIKPAIEMNGSDPLCTEMQASN 1080
DB 5236 EALKNKSQNPELD-----KVTIE-----QAIN 5257
QY 1081 SVPTSVHQLRPADIKVVAALGDSLTTAVGARPNNSDDLPTSMRGLSMSTIGDNLERTHTT 1140
DB 5258 NINSAINNLN-GEAKLTAKEDAV-----ASINLISGL----- 5289
QY 1141 LPLNLIKFPYLLGFTSTWEGTAGLVNVAEGARAPMAQAW-----DLV 1186
DB 5290 -----TNEQKTEKNOAVGAQTRQVANKLDAEALDQSMQTLRDLV 5331
QY 1187 ER-----MKKSPINIEKQKMLVTLPIGVNDLCHYCEBNPAHLAT--EX 1228
DB 5332 NNQONAIHSTSNYFNEEDSTOKNTYDMAILDNGSTYIT--GQH-----NPLINKSTIDQT 5381
QY 1229 VOHIQALDILS--EELPR-----APVNVVWELASLYOQGGKCMILAON 1274
DB 5382 ISRTINTAKNDLHGVKEKQORDKGTANOELGOLGYLNDPOKSGEESLVNGSNTRSEVEEHLN 5441
QY 1275 NCTCLRHSGSLKQEKELKVNNVNLQHGISFSPYHQ--YTQREDFAVAVVOFFPONTLTPL 1332
DB 5442 EAKSINNAMQOLRAKVAEKTVKQSSDYINDSTEHQGYDA-----LOEAMNII 5491
QY 1333 NERGDUTL--TFSEDDCFHSSDRGHAEALALM---NNMLPEVGRKITSNNFTHSRA 1384
DB 5492 NEIGNPTLNKSEIEQKLOQLTDAQNALQSHLEAKKNNAITGINKLTALDADQROKA 5549

```

```

RESULT 2
US-08-726-012B-2
; Sequence 2, Application US/08726012B
; Patent No. 5952190
; GENERAL INFORMATION:
; APPLICANT: Hans Joenje, et al.
; TITLE OF INVENTION: CDNA FOR PANCONI ANEMIA COMPLEMENTATION GROUP A
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klaarquist Sparkman Campbell Leigh & Whinston, LLP
; STREET: One World Trade Center, Suite 1600, 121 S.W. Salmon Street
; CITY: Portland
; STATE: OR
; COUNTRY: USA
; ZIP: 97204-2988
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Disk, 3.5-inch
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: MS DOS
; SOFTWARE: Mordperfect 5.1+, ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726, 012B
; FILING DATE: 10/04/96
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Richard J. Polley
; REGISTRATION NUMBER: 28,107
; REFERENCE/DOCKET NUMBER: 3812-45520/RJP/DJB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (503) 226-7391
; TELEFAX: (503) 228-9446
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1455
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: Linear
; US-08-726-012B-2

Query Match 1.8%; Score 136; DB 2; Length 1455;
Best Local Similarity 19.3%; Pred. No. 0.0032;
Matches 225; Conservative 133; Mismatches 393; Indels 408; Gaps 58;

QY 143 IOAQELVRNKNENIOLDFQFQWKLINVFPNSAQCYLCPESAQONGLAAGVDELMG----- 198
DB 366 ISABELVGHQVLELQ-EVHWQVLSFVSALVVCF--PEAQO--LLEDVVARLMAQAFE 420
QY 199 -----VLDYIQQEVPAFVNLVLDSEVAEVS-RQYHGTWLSAPPECNGSEET 245
DB 421 SCOLDSWTAFLVVRQALLEGPSAFLSYADWFASRGSTGYHG-----CSKKA 469
QY 246 -----TFLAKVNWQMSVQEAAMNSLLASRYSEQESFTVVPQPFYET--TPSLHSEDPRL 298
DB 470 LVLFLEFLSELV-----FESPRYLQVHILHPLVPEKYSLSLTDYISLAKTRL 518
QY 299 QDSTTLAMHLMN-----RMMEPAGEKDEPLS-----VKH-GRPMKCPQESPYL 341
DB 519 AD--LKVSIENMGVLEDLSASAGDITEPHSQALQDVEKAIMVEPHGNINIVTVWMAISIFR 575
QY 342 FSTRNSNYLTRLOKPODKLEVRREGAEIRCPDPSDTPVPSVYRL-----KRADI 391
DB 576 RPYVYSHFLPALLTPRLPVPDPSRVAFISLKRADKIPPSLSTVYCOACSAAEKPED- 634
QY 392 NVIGALGDSLTAAGAGSTPGNVLDVLTQYRGWSVSGGENTGVTTLANILREFPRL 451
DB 635 -----AALGVRAEENIS-----AEEPQQLTALGEIR----- 661
QY 452 KGFSVGTGKETSPN-----AFINQ--AVAGRAEDLPVQARLVLMKNQDTRIH 499
DB 662 -----ASMTDPSQKDVISQAQVAIVISERLRAVLGNHEDDSSVEISK-IOLSTINTPLRP 713

```

```

QY 500 QEDMKITITLFIGNDLCDPCNDLVHSPONFTNIGKALDILHAEPRAVNLVTLEIV 559
DB 714 REHIAVDL-----LTSFCQNLMAAS-----SVAP-----738
QY 560 NRELYOEKKVYPRNIIKSLC-----PCVLFKPDNDSTELATLIEFNKQOEKTHQ-----611
DB 739 -----PERGFWAALFVTRTCGRVPAVL-----TRLCQLRHQGPSLSAPHVGLAA 786
QY 612 ---IESGRYDREDFTVVQPFENVMKPSBGLP-----DNSFPADCF- 654
DB 787 LAVHLG--BSRALPREV-----DVGPAPGAGLPVPAFLDSILTCRTRSLF-----CLK 835
QY 655 -----HFSKSHSRAASALWMNMLEPVGQKTRHKFNKINITCPNOVPLRT 703
DB 836 FCTAASYSLSCKSSQSDRLCCLSPGLIK-----KFOPLFRLESEKAPOLSEE 886
QY 704 YKNSMOGHGTWLPDRAPASAL-----HPTSVAHLRPADIQVVALGDSLTAAGIGSKPD 759
DB 887 DVASLSMRPLHPSADMQRALSLMTHRTREVLKEEDVHL--TYQDWLHLELEIQEAD 944
QY 760 DLBDVTQ-----YKGLSYAGSGSLENT--TLPILIREFNRLTGAVGTGD 807
DB 945 ALSDTERODFHQWAIHFNHLPRESSASGCGDGLQAACTTLVNALMDFHSGRSY-----D 999
QY 808 ANDTNAFLNOAVPGAKA--EDLMSOVQTLMOKKMDHRAVNFHEDMKVITVLIGSDLCQY 865
DB 1000 HSEN---SDLVFGGTGNEIDITSRLQENVADE-----LQOD--LIVPLGHTP--- 1042
QY 866 CTDNSNYSANFV--DHLRNALDVL-----HREV--PRVLVNLVDLNFPTIMR 909
DB 1043 -----SQEHFLFBJFRRLQALTSGMSVAASLQGRMLMKRILRL----- 1085
QY 910 QVFLGPRDKPCVQOASVLCVLTLENSQELARLEAFSAVSSNRELVGSGRYTQED 969
DB 1086 -----PSSVLCGS--SFQAEORITARCQFFHLVNSEKRNCFSHGALTOD- 1129
QY 970 FSVVLQPFONIQPLVADGLPD--TSFPARDCHNPKHQSOLADAL--WTNMLPLPGSK 1026
DB 1130 ---ITHFRGLNACLRSRDBSLMVDFLAKC---OTKCPRLTSLALVMPSPLEPV--- 1180
QY 1027 TETLDLRAEMPTICPTQNEBFLRTPRNSNYTPRIKPAIE--NMWSPFLCTE----- 1075
DB 1181 -----LTCRRRHQGSPLPRE-----LQKQEGRGFASDPLSEASAPANPD 1223
QY 1076 WKASNSVPTSVHOLRPADIK-----VVAALGDSITTAVGARPN 1114
DB 1224 WLSAALHFAIQVREENIRKQKLDCEEREBLLVFLFFSLMGLLSHLTS-----NS 1277
QY 1115 SSDLPTSM-----RGLSM 1127
DB 1278 TTDLPKAFVCAALIECLEKRIKISM 1302

RESULT 3
US-09-589-567-2
; Sequence 2, Application US/09589567
; Patent No. 6479730
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; APPLICANT: Shi, Jinru
; TITLE OF INVENTION: Maize DNA Ligase II Orthologue and Uses
; FILE REFERENCE: 1125
; CURRENT APPLICATION NUMBER: US/09/589,567
; CURRENT FILING DATE: 2000-06-07
; PRIOR APPLICATION NUMBER: US 60/145,911
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1218
; TYPE: PRT

```

```

; ORGANISM: Zea Mays
; US-09-589-567-2
Query Match 1.7%; Score 129.5; DB 4; Length 1218;
Best Local Similarity 18.6%; Pred. No. 0.01;
Matches 237; Conservative 151; Mismatches 418; Indels 471; Gaps 60;

QY 10 LELLILGQGTPOIHTSPKSTLEGQIMPELTKNSPP-----CNPKLVNMPKSVHS 64
DB 90 MKILTLGFGG-----KNSVFTEDACTDTHVGLMNLGSETSPY 128
QY 65 LKPSDIK-----FVAIG-----NLEIPDPGTGLEKOD 94
DB 129 FQNFVFKLEIMMERGAKAVGFVPTGMVYETKKEGFAVRXDSLEIHLVPS---EHS 185
QY 95 WTERPQVCMGVMTVLSDIIRFSPSPVPVCHTGKRVYRPHQAEMLQAOGLVNNKKE 154
DB 186 YNE-----LRDYVKFLHP-----KSVIFTAGVDAGKLDKSKAFAV--- 219
QY 155 NQLDLPQFDMKLIINVFPSNASOCYLCPSAQONGLAAGVDLMDGVLDYLOQEVPRAFVNL 214
DB 220 LQKHF-----AGLVDETANKQEFMAFHRS--- 245
QY 215 VDLSE-----VAVSRQYHGTWLSPAPEPCNGSEETTRIAKVWQ-----WSY 257
DB 246 IATRECKDVVNCSSQLDGEDALLPAISASEQDCKLEMENTQETLKEISDFLPWSYS 305
QY 258 QEAMNSLASKSYSEQESFTVVQ--PFFYETPSLHSEDPRLQDSTTLAMHLNRIME 314
DB 306 QEOITDILLMSSEDDVYKAKASIFFERERDFEEANVSCNG-----T 345
QY 315 PGEKDEPLSVNGRPMKPSQESPYLFYSRNSNYTRL-----QKQP--DKLEVRG 365
DB 346 PTEKKNH--SSNHGSADVSSQOESLSQKVEYSAKLVNISPMMKRPVTPKVKNGGS 403
QY 366 AEIRCPDKPDSPTVP-----SVARLKPADINVI GALDSITAGN 405
DB 404 SMTNPKKKGRILSTSESGRKOSTITNYFIRATAFAFKRCRPERVTEAHQNVENSQ 463
QY 406 GAGS--TPGNVLDVLTQYGLSNVSG-----DENIGTVTTLANILREPNLSK 452
DB 464 LADTEKTOQOIVDOLQI-----VDGSMREYAVYLIEKAGDVAVAIDI---FYSSSE 514
QY 453 GSVGTGKETSNAFLNQ-----VAGGAEDELVQOARLVLDLMDKDTIHHQ 500
DB 515 GNNV---IEVDKNIMQNAQGETADNCSETDMASGVQATP----- 552
QY 501 EDKKITITLFIGNDLCDPCNDLVHSPONFTNIGKALDILH---AEVPRAFVNLVTVL 556
DB 553 ---KMSNLVQ-----TSLAHGDSANTSLIEKYLEIEHACWIAQGPARYLHARTF 601
QY 557 EIVNIRELYOEK---KYVC-----PRMILSLCPVLYK---DNDSTEL--- 594
DB 602 DIVE-----KEKGIKITAIFCNMFRSLALSPDDVLPVAVLCTNKISPDHEMELNIG 656
QY 595 ---ATLIEFNKKFQKTHQILBS--GRVYREDFTVVVOPRENV--DMKTSBGLPDNSF 648
DB 657 SLVVTTLLESIGTSRKIHMYKTYGDLRCRQNTLAPPSLSIDVSTJAKL----- 711
QY 649 PARDCFHFSKSHSRAASALWMNMLEPVGQKTRHKFNKINITCPNOVPLRTYKNSM 708
DB 712 -----SAISGSGAG-----RRK-----NLVHLIRS----- 733
QY 709 QGHGTWLPDRAPASALHPTSVHALRPADIQ--VVAALGDSITAGNGIGSKPD--DLPDVT 765
DB 734 -----CRETEMKFLVTVLRNLRIGAMMKTIIPALAHAVGFDRCNADRAVSLGK 785
QY 766 TOYRGISYAGDGSLENT--TLPNIIAREFNRLTGAVGTGDADDTNAFLNOAVP 820
DB 786 SOLQSLSTEVE--AEAYVTPLNLDLIPSLRE-----GTAFSSSLAM---VP 828
QY 821 GAKAEDLMSQVOT--LMQKKMDHRAVNFHEDMKVITVLIGSDLCQDCTSNYSANFVD 879

```


Db 829 GTPIPMRLARITNGLTOSLKAH-----GRSFCEYKDG---QRAQIHK 870
 QY 880 HLRNALDVLRREVRLVNLVDPLNPTIMROVFLGNPDKCPVOOASVLCNC--VLTREN 937
 Db 871 ALGGSVIRFSRQKESIRPPDLVN--MIKEL-----CKEYSSITLDAEVGINRKK 921
 QY 938 SOELARLEAFSAVRSSMRLEVSGRYDQEDPSVVLQPFQNIQLPVL-----DG 989
 Db 922 GNEI-----MSFQSLSSRERGNKHSISI-----ENIKVDICVFVDIMFHG 964
 QY 990 -----LPTSPFAPPCITIPNOKFHSQALRALMTNMLEPLGSKTEITLDAEMPT---- 1038
 Db 965 ERLDCEPLRQRKRYIHDLFEKPGYF-ELAQOLNIEACEASPDNSSTL---ARMDTFERN 1020
 QY 1039 TCPQNEPFLRTPRNSVYVPIKPAIENMGSDFLCTEWMKASNSVPSVHQLRADIVY-- 1096
 Db 1021 ACSSCEGIMKTLVDVAGYSASAKRCDSW-----LKVKR 1054
 QY 1097 --VAALGDSL-TTAVGA 1110
 Db 1055 DYVGAGDSLVLVPGA 1071

RESULT 4
 US-09-180-439-8
 ; Sequence 8, Application US/09180439
 ; Patent No. 6225532
 ; GENERAL INFORMATION:
 ; APPLICANT: Dixon, Mark S
 ; APPLICANT: Hatzixanthis, Kostas
 ; APPLICANT: Jones, David A
 ; APPLICANT: Jones, Jonathan DG
 ; TITLE OF INVENTION: Plant pathogen resistance genes and uses thereof
 ; FILE REFERENCE: 620 - 53
 ; CURRENT APPLICATION NUMBER: US/09/180, 439
 ; EARLIER FILING DATE: 1998-12-06
 ; EARLIER APPLICATION NUMBER: PCT/GB97/01249
 ; EARLIER FILING DATE: 1997-05-08
 ; EARLIER APPLICATION NUMBER: GB 9609681.3
 ; EARLIER FILING DATE: 1996-05-09
 ; EARLIER APPLICATION NUMBER: GB 9619924.5
 ; EARLIER FILING DATE: 1996-09-24
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: Patent Ver. 2.0
 ; SEQ ID NO 8
 ; LENGTH: 1016
 ; TYPE: PRT
 ; ORGANISM: Lycopersicon esculentum
 ; US-09-180-439-8

Query Match 1.64; Score 127.5; DB 3; Length 1016;
 Best Local Similarity 19.64; Pred. No. 0.011;
 Matches 207; Conservative 134; Mismatches 380; Indels 333; Gaps 55;
 QY 40 TLKNSPPCPNKLGVMPSPKSVSLKPSDIKFVAAGNT-----EIPPD 84
 Db 85 TLVAFPPSSLPLENDLSNNNISGTPPE-----IGNLTNLYVDLNTNNOISGTPPQ 138
 QY 85 PGTGDLKQDWT-----RPQYV-CMGVMTVLSDIIRYSPSPVPMVCHTGKRVIP 134
 Db 139 --IGSLAKQIIRIFNNHLNGFIPBEIGYLRSLTKSLGINFLSGSIPASIGNM----- 190
 QY 135 HDGAEDMIAQOELVRRMK-----NLQDDFQ-----DWKLINVEFSNASQCYLC 180
 Db 191 -TNLSFLEVENOISGFIPEIGYLRSLTKSLDINLSGSSIPASLGNL--NNLSPLYL- 246
 QY 181 PSAQONGLAAGVDELGVL-----DYLOQEVPRAFVNLVDSEVAEVSROYHGT 230
 Db 247 ----YNNOLSGSIPBEIGYLRSLTKSLGINFLSGSIPASIGNLNLSRLDLYNNKLSGS 302
 QY 231 WLSPAPPCPCSEETTLAKVVMQWSYQEAAMNSILASSRYSSEDSFTVVFQPFY--ETT 288
 Db 303 ----IPEIGYLRSLT-----YLDGNNALNGSIPASIGNLNLSRLDLYNNQLS 348

QY 289 PSLHSEDPRLQDSTTLAMHLMNRMMEPAGEKDEPLSVKIGRPMKCPSESPLYFSYRNS 348
 Db 349 GSIPBEIGYLRSLTYL-----DIGNMANGSI-----PA-----SLGNLN 383
 QY 349 YLTRLOKPODKLE-----VRGAEIRCPDKSDPVTVPYSVRLK----- 387
 Db 384 NLSRLDLNKKLSGSIPBEIGYLRSLTYLDLGENALNGSIPASIGNLNLSRLDLYNNQL 443
 QY 388 ----PADIVIGAGLSLTAAGNCA--GSTR-----GNLVDVLTQY---RGLSMWSGDEMG 435
 Db 444 SGSIPEIGYLSLTYL-VYGNNSNGSIPASIGNLNLSRLDLYNNQLSGSI--PEIG 500
 QY 436 TTTTLNIIREFPSSIKG---FSVGTGKETSPAPFLNOAVAGGRADLPQARLYDLMK 492
 Db 501 YLSLTYELFG-NNSNGSIPASIGNLNLSRLDLYNNQLSGSI--SIPASFGNNMLQ- 554
 QY 493 NDTRIHFQEDWKIYLFIGNDLCD----FCN---DIVHSPQNTDNIGAL----- 538
 Db 555 -----TLFLSDNDLIGEPISFVCNLTSLVLYMSRNNLKGVPOCLGNISD 600
 QY 539 -----DLHAEVRAVNLVTVLEIV-----NLREL-----YQEK- 569
 Db 601 LHLMSNSMSPRGELPSSISNL-TSLKIIDFGNNLEGALPOFGNISLQVFDMMNNKL 659
 QY 570 -----VYCPRMIL-----RSLCPC-----VLKFDNSTE-----LATLI 598
 Db 660 SGTLPNPFISGGLISLNLHGNELADEIRSLDNCCKQLVDLGDVQLDTPFMWLTLP 719
 QY 599 EFN--KKFOEKTHQLESGRYDREDFTVVQVFPFENVMPKT--SEGPDNSFPAPDCE 654
 Db 720 ELRLVRLTSNKLHGPIRSSGAE-----IMPDRLIIDLNRNAFQDIPSLF----- 766
 QY 655 HESSKSHSRAASLYMNMLEPVQOKTRKFEKINITEPN--QVQPLRTY-----K 705
 Db 767 -----EHLGMRVTQDTEEP---SYESYDSDVVVTKGLELELVRLSLTYIIDLS 817
 QY 706 NSMOGHGTWLCRDRAPSAHLPTSVHARLPADIQ-----VVAALGD-----SLTAG 751
 Db 818 NKREGH-----IPSVL--GDLAIRLVNLSHALOGYIPSSIGLSILESDDLSTN 866
 QY 752 NGISGRDDLPDVT-QYRGLS--YSAGDGSLENTYTLNLIREFNRNLGYAVGTGDA 808
 Db 867 QLSGEIPQOLASLTFLVNLNSHNYLQGCIPQGPQRTFESNSYEGNDGIRGYPVSKGC 926
 QY 809 NDTNAFLNOAVPQAKEDLMSQVQITLMQCKKDDHRVNFHEMDKVTIVVLG-GSDLC- 863
 Db 927 KQVSEKNYTVSALDEQESNSEP-----FNDPMK--AALMGYSGLCIGIS 970
 QY 864 ---DYCTDSNLYSAANFVDHLNRALDVLREVR 894
 Db 971 IYIYLSTGNLRLARIIELEHKITVQRRKOR 1004

RESULT 5
 US-09-252-991A-25753
 ; Sequence 25753, Application US/09252991A
 ; Patent No. 6531795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196, 136
 ; CURRENT APPLICATION NUMBER: US/09/252, 991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074, 788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094, 190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 25753
 ; LENGTH: 2636
 ; TYPE: PRT

ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25753

Query Match 1.6%; Score 123.5; DB 4; Length 2636;

Best Local Similarity 21.5%; Pred. No. 0.2; Mismatches 264; Indels 215; Gaps 39;

Matches 154; Conservative 83; Mismatches 264; Indels 215; Gaps 39;

814 FLNQAIVGAK-----AEDLMGVOITLMQKMDHHRVNFH---DMKV--I 853
1121 FFEQPFPPNQHNMOSLLKREALKALALQALVEH-HDMLRLRFHETDGTMAEHA 1179
854 TVLIGSSDLCDCYCTDSNLVSANFVDHLRNALDVLHREVRVLVNLVDFLNPIMEQVFL 913
1180 EATLGA-----LWRAEAVD--RQALSLCESSQSL---DLADGLRLSLIV 1223
914 GNDKCPVOQASVLCNCVLTRENSQELAEFSRAYSMSR---ELVG-----S 961
1224 ---DMADGGORLLVTHLVDVSWRL-LEDLQRAYQOSLGEAPRLPGKTSPPKAWA 1279
962 GRVDTQ---EDFVVLQPFQNIQLPYLADGLPDTSPFADCIHP---NQKF---HSQ 1010
1280 GRVSEHARGESMAQLO-FWRE---LLEGAPAEI---PCVHPQALEQRFALISVQSR 1329
1011 LARALWTNMLE-PLGSKTETLD-LRAEMPITCPTONEBPFLTPRNSVYPIKPAIEN 1066
1330 FDRSLTERLLKQAPAYRTQVNDLLTLALRVV-----RMSGASSL-VQLEG 1377
1067 WGSDFCTEKKANSVP--TSVQLRPADIKVVAALGDSILT--AVGARPNNSDLPTS 1121
1378 HGRSEELFADIDLSRTVGWFTS--LPPVRLSPVADIGESIKAIKEQLRAIPD----- 1426
1122 WRGLSMS---IGDDNLEHTHTLPNLIKKNPYLLGFSTWEGTAGLVNAEGARARD 1177
1427 -KGLGYGLRLYLGESARVLAGLPQARITFN--YLGQPDQDPDWMALDPAGESAGAEM 1483
1178 MPQAWDLVERMKNSPDIN-----LEKDWKLVTLFTIGVNDLCHYCENPEALATEVYQ 1230
1484 DECAPLD-----NWLSTNGRVDFGELSTIDMSFSQMGF-----EDQVRLADLYVA 1529
1231 HIOQALDIISSE-----LPPAFVNVVELASLYGGGQKCA 1268
1530 ELTALDFCCDSFRHGATSPDFLAGLDQARLDALPVALEVEDIYPLSPMOG----- 1583
1269 ML-----AAQNNCTCLHNSOSSLEKQELKYNMNL---QHGISFSYHQTQREDPA 1318
1584 MEFHSLYEQASSDYINQMRYDVSGLDI PRFRAA-WQSALDRHALILSGFAMQGLQOPIQ 1642
1319 VVVQPFQNTLTPINERDITLTFESDECFHFSDRGAEWALAMNNMLEPVRKTTSNM 1378
1643 IYVR-----QRLPFAEEDLSQAANRDAALLALA-----AAREBG 1678
1379 FTHSRAKL-----KCPSPSPYLYTLRN-----SRLLPDQAEAEV 1416
1679 FELQAPLRLRLLVKTAEGEHLIYTHHLLDGMGNAQLLSVLSVAGRSPEQL 1734

RESULT 6

US-09-107-532A-5646
Sequence 5646, Application US/09107532A
Patent No. 6583275

GENERAL INFORMATION:

APPLICANT: Lynn A Doucette-Stamm and David Bush

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSER: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354

COMPUTER READABLE FORM:

MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781) 893-8277
TELEFAX: (781) 893-8277
INFORMATION FOR SEQ ID NO: 5646:
SEQUENCE CHARACTERISTICS:
LENGTH: 855 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...855
SEQUENCE DESCRIPTION: SEQ ID NO: 5646:
US-09-107-532A-5646

Query Match 1.6%; Score 122; DB 4; Length 855;
Best Local Similarity 19.4%; Pred. No. 0.029;
Matches 134; Conservative 88; Mismatches 237; Indels 230; Gaps 36;

357 QDKLEVBEGAEIRCPKD--PSDTVPTSVRLKPADINIGA---LGSLTAGNAGST 410
118 QEAVLVSNGMQEIKPKELVKGDIYVLRAGDVVPADIRIEAHDLVSEISILGSESEASE 177
411 PGNVLDLVNRYGLMSVSGDENGIVTTLNLRFRNSLGSFSGTQKERSPNMFLNQ 470
178 KSHA--VLTEER---SIDQKNMGSGTLVQ-----SGSAGVVAVETANTTEIGK-INV 225
471 AVAGRAEDLPQARRLVLMKN-----DRIHFOEDWKITTLFIGNDLC 516
226 ALQSVQQTPL-IKINRLNQIFQGLICLSLFLVITFTFRYGMNHIL----- 274
517 DPCNDLVHYSFONFTDNIGKALDIHAEPRAFPVNLVTVLEIVNLELYQEKKVY--CPR 574
275 -----LSATTALVVSMPREGLPALVTILSVGVHEMAKEXAIIKGLPS 317
575 M-----ILRSLC-----PCVLKEDDNSTELATLIEPKKROEKT 608
318 VETLSMTVICSDDKTGLTKNEMTVYVDAKEBACVLSIMKQCQELKT-----KEQKT 371
609 HOLIEGRYDREDFTVVVQPFENVDMPTSEGLPDNSFF----- 649
372 ENL--SG---NPEVALIHETETANLPLKIKIAKIPSSSEYKMTMAHEKKAIVYVG 426
650 ADPCFPHSSKSHRASAALMNNMLEBVGQKTR-----HKFENKINITCPNOVQPLRT 703
427 ABEVL-FAKSTLSAEQEMSQTAAEFAKRGQVLCFAVKYVDSKQELTHELTLO--LT 482
704 YKNSMGCHTWMPCGDRASALAHPTSVYHLRPAADIOVAALDLSLTAAGSIGSKPPDLD 763
483 FAGIA--GLIDPPKSAVKA-----VKECQAGISVKMITGDH-----KDTAK 523
764 VTTQYRGLSYSA-----GQDGSU-----ENVTTLP-----NLRREFNRNLATGY 801
524 AIAEOIGLKHAKVLEGIDLDLMSDEBLIQVPYIVDFVARTTPPHKRLRYKALQKN--GE 581

QY 802 AVG-TGD-ANDTNAFLNQA-V--PGAKABEDJMSQVOTLMOKMKDHRVNEHEDMKVITVL 856
DB 582 IYVMTGDGVVADAPALKRSVDGIAMGIGSEVSQAAMV--LGD--NHTTAKAV--- 633
QY 857 IGSGLCDVCTDSNLYSANFVDFHRLNALVLRHREPRVLVNL--VDFLNPITMRQ--- 910
DB 634 -----KEGRRIIDNLQKTIINFLPTALAQGLI 660
QY 911 ----VFLGNPKCPVQQA-VL-CNCVLT.L 934
DB 661 LIALMLNRP-LPLSPVQILMVMVTTI 687
RESULT 7
US-09-353-585-2
Sequence 2, Application US/09353585
Patent No. 6287865
GENERAL INFORMATION:
APPLICANT: Dixon, Mark S
Jones, David A
TITLE OF INVENTION: Plant pathogen resistance genes and uses
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESS: 8th Floor, 1100 No. 6287865th Glebe Road
CITY: Arlington
STATE: Virginia
COUNTRY: United States of America
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/353,585
FILING DATE: 15-Jul-1999
CLASSIFICATION: C12N 15/29, 15/82, A01H 5/00, A01N 65/00, C12Q 1/68
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/930,277
FILING DATE: 27-OCT-1997
APPLICATION NUMBER: PCT/GB96/00785
FILING DATE: 01-APR-1996
APPLICATION NUMBER: GB 9506658.5
FILING DATE: 31-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Ms Mary J Wilson
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 620-69
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1112 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Tomato
STRAIN: Cf2
US-09-353-585-2
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
Query Match 1.6%; Score 121.5; DB 3; Length 1112;
Best Local Similarity 18.2%; Pred. No.: 0.055;
Matches 196; Conservative 139; Mismatches 356; Indels 387; Gaps 48;

QY 53 LGVNMPSKSV-----HSIK-----PEDIKFAAIGNLEIPDPGTDLEK 92
DB 174 LGINIFSGSIPASVGNLNNISFLYLNNQSGSIPBEISVLRSLTDLSDNLANGSI-- 231
QY 93 QDWTERPOOVCMGVTVLSDIIRY---FSPSVMPVCHGKRVIPHDGAE----- 140
DB 232 -----PASIGNNNLSFLFYGNQSGSIPBEICYL-RSLTYLDSLSENLANGSIPAS 282
QY 141 -----LMLQAQL-----VNMKE----- 154
DB 283 LGNINLSFLFLYGNQSGSIPBEIGYLSNLVGLSENLANGSIPASGLNKLNSRLNL 342
QY 155 -NLQDFQFDMKLINFFSNASQCYLCPSAQNGLAGVDEIMGVLDYLOQ----- 205
DB 343 VNNQSGSIPASIGNL--NNLSMLYL-----YNNQSGSIPASIGNLNNLSMLYLNNQL 395
QY 206 -EVPAPVNLVLDSEVAEVSROYHGTWLSAPAPCNCSEETRLAKVVMQMSYQEAAMS 263
DB 396 SGSIIPASIGNLNNLSRLYLNNQSG-----SIPBEIGYLSLTYLDSLNNNSING 445
QY 264 LLAAS-----RYSRQESFTVVPQFPFYETTPSLHSEDPRLQDSTTLAMHLNRM 312
DB 446 FIPASFGNNSNLAFLFLYENQLASVPEBIGYLSNLVDSLSENLANGSIPASIGNLNL 505
QY 313 MEPAGEKDEPLSYHGKPMKCPQESPYL-----FSYRNSNYLTRLO 354
DB 506 SR-----LNLVNNQSGSIPBEIGYLSNLVDSLSENLANGSIPASIGNLNNLSRLN 557
QY 355 KPQDKLEVBEGABEIRCPDPDPTVPTSVHRLKPADINIVIGALGDSLTAQNGAGSTPG-- 412
DB 558 LVNNQL-----SGSIPBEIGYLR--SLNDGLSENLN-----GSIPASL 595
QY 413 ---NLVDVLTQYRG-LSMSVSGDENIGTYTTLANILREPRPSLKGVSVCGETSPAPL 468
DB 596 GNINLSMLYLVNNQSGSI--PEEIGYLSL--TYSLGNSLNGL-----IPASFA 644
QY 469 NQAVAGRAEDLPVQARRVLDLMMKNDTRIHFQEDMKITLFIQ--GNDLCPGNDLVHYS 526
DB 645 NM-----RNLQALILNDNNL-----IGSISVSVCNLSLEVLVM 678
QY 527 PON-----FTDNIGKAL-----DILHAEPRA 548
DB 679 PRNNLKGVPOQCGNISNLQVLSMSSNSFEGELPSSISNLTSQILDGFRNNLEGAIPQC 728
QY 549 FVNLVTVLEIVNRE-----LYQEKVYCPRAIL-----RSLCPG-----YLK 586
DB 739 FGN-LSLEVFDMQNNKLSGTLPTNFSIGCSLISLHNGELEDEIPLSLDNCKKLQVLD 797
QY 587 FDNSTG-----LATLIEFN--KKFOEKTHQLESGRVYTRBDFTVVQPFENVDMPK 638
DB 798 LGNQLNDTRPMMGLTLPBLRVLRSLSNKLHGIRSRAB-----IMPDRLIIDLSR 850
QY 639 T--SEGLPONSFPAPDFHSSKSHSRAASALNNMLLEPVQK----- 679
DB 851 NAFSODLPSTLF-----EHLKCMKRTVKTMEPESEYDSDVVVTKGLELEI 899
QY 680 -----TRHAKFENKINITCPNQVQ--PLRTKYKSMQGHGTLPCRDAPSLHP 727
DB 900 VRILSLYTVLDSNNKFEHGISVLDGLAIRILNYSNALQGY-----IPSLGS 950
QY 728 TSVHALRPADIQVVALGDSLFTAGNGIGSKPDDLPTVT--TOYGLS--YSAGDGSLENV 784
DB 951 LSI--LESID-----LSFNQSGSIPQOLASITFLFNLNLSNLYQGGCIPQGPOR 998
QY 785 TTLPNILREFFRNRLTGAVAGTGANDTNALNQAIVGAKAEDJMSQVOTLMOKMKDHRV 844
DB 999 RTFESNSYEGNDGLRGVPSYSGKGPVSEKNTVTALEQESNSER----- 1045
QY 845 NFHEDMKVITVLIG-GSDIC-----DYCTDSNLYSANFVDFHRLNALVLRHREPR 894
DB 1046 -FNDFWK--AALMGYSGLCTIGSMIYIILSTGNLRLWLAIRIEKLHKKIMQRRKQR 1100

RESULT 8

US-07-741-940-7
Sequence 7, Application US/07741940

Patent No. 5352775
GENERAL INFORMATION:

APPLICANT: ALBERTSEN, HANS
APPLICANT: ANAND, RAKESH
APPLICANT: CARLSON, MARY
APPLICANT: GRODEN, JOANNA
APPLICANT: HEDGE, PHILIP J.
APPLICANT: JOSLYN, GEOF.
APPLICANT: KINZLER, KENNETH
APPLICANT: MARKHAM, ALEXANDER F.
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: THILVERIS, ANDREW
TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie & Beckett
STREET: 1001 G Street, NW
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001-4598

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/741,940
FILING DATE: 19920109
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 1107.035574
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9299
TELEFAX: 202-508-9100

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:
LENGTH: 2842 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: APC
US-07-741-940-7

Query Match 1.6%; Score 120.5; DB 1; Length 2842;
Best Local Similarity 19.8%; Pred. No. 0.46;
Matches 184; Conservative 104; Mismatches 328; Indels 311; Gaps 47;

QY 34 GQLMPELTLSKPPCPNPKLVGMVPSKSVSLKPSDIKFAVLAIGNLEIPDPGTGDLK-92
DB 1920 GQKPFILQKOSTP-----QSSKDI-----PDRKAADK 1950
QY 93 QDWTERPQOVCV--GVMTVLSIIIR-----YESPSVPM-----123
DB 1951 QNFALNTPVCFSHNSLSLSIDIDENNNKENEPKETEPDPSQGPSPKQASGYAPKS 2010
QY 124 -----PVC--HTGKRVIPHDGADLMIAQELV-----RNM 152
DB 2011 FHVEDTPVCFSSRSSLSISIDSEDDL--LQECISSAMPKKKPSFLKGDNEKHSFRNM 2067
QY 153 -----KENTLQDPQ-----FDWKLII-----NVFFSNASQ--CYLCPESAQ 185

DB 2068 GGILGEDLTLDKDIORPDESHGLSPDSENFDMKALIOEGANSIVSLHQAAAAACLSRQ- 2126

QY 186 NGLAGVDELMGVLDVYLQOEVPRAFVNLVDLSVAEVSQHYGTWLSPPAPPCNCEET 245

DB 2127 ---ASSDSDSLTKSGISLGGSP--PHLLPDQEKPEPTSNK-----GPRLLKPEKS 2173

QY 246 TRLAKVVMQWSY-----QEAAMSLASRSYSEOSFTVVQPPFETTPSLHSEDPRLQD 300

DB 2174 TLETKIESSEKIKKGGKYKSLITGKVRNSNEISGMQMP-----LQNMPSISR 2225

QY 301 STLLAWHLNRMMEPAKGEDEPLSVHGRPMKCPDSQSPYLFYSRNSNYLTRLOKPODKL 360

DB 2226 GRIMI-HIPG--VRNSSSSPVS-KKGPLKTPASKP--SEGQATTPSPRGAKPSVKS 2279

QY 361 EV-----REGAEIRCPDKPS-----DTPPSVRLKPADINVIGALGDSLTAAGAGSTP 411

DB 2280 ELSPVARQTSQIGSSSKAPSSGRSSTPBRPAQ--QPLSRPIQSGRNSISPGNGISPP 2338

QY 412 GNVLDVLTQYRGLSWGVGDENIGVTTLANILREFPNLSLKG--SVGTGK--ETSPPAFL 468

DB 2339 NK-----LSQLPRTSSPSTASTKSSGSGCMYSTSPGRQM 2372

QY 469 NQAVAGRAEDLPVQARRLVLDLMDNTRIHFOED-WKITLFTGNDLCDPCNDLVHYS 527

DB 2373 SQ-----QNLTKQ--TGLSKNASIPRSBSASGKLQNMNG--ANKVELSR 2417

QY 528 QNTDNIGKRLDLIAEVPRAFVNLVTLVLEIVNLRELQYCKKYCPMILRSCLPCVLKF 587

DB 2418 MSTKSSGSESD--RSERP-----VLVROSTRKE-----ASPILR-----2452

QY 588 DNMSTELATLIEFNKFOEKTQILIESGRYDTEDFTVVQPPFENVDMKTEGLPDNS 647

DB 2453 -----KLEESASGESLSPSRPASPSRQAQTVLSPSLPD-- 2489

QY 648 FPAPDFHSSKSHSPASALMNM--LEPV-----GQKTRHKKENKINITCPNOVOP 699

DB 2490 -----MSLSTHSSVQGVGRKLPNLSPTLIEYNDGRPAKHDLA--RHSSPSRL-P 2539

QY 700 FLATYKNSMOGHTVLPCCRPAASALHPTSVAHLRPADIOVVAALGDSLTAAGISGKPD 759

DB 2540 INRS-----GTWREHSHKSSSLPRVSTWRTGSSSSILASSESSER-----AKSE 2586

QY 760 DLPDVTQYRGLSYSGDGLSENVTLTPILIR-----EPN--NLTGYAVGTGDANDTNA 813

DB 2587 DEGHVNS-----IGTKOSKENOVSAKGTWAKIKENEPNSISQTVSSGATN-----2635

QY 814 FLNQAVGAKAEDLMSQVOTLMQKMD 840

DB 2636 -----GAEKTLTIQMAPAVKTED 2655

RESULT 9

US-08-289-548A-7
Sequence 7, Application US/08289548A
Patent No. 5648212

GENERAL INFORMATION:

APPLICANT: ALBERTSEN, HANS
APPLICANT: ANAND, RAKESH
APPLICANT: CARLSON, MARY
APPLICANT: GRODEN, JOANNA
APPLICANT: HEDGE, PHILIP J.
APPLICANT: JOSLYN, GEOF.
APPLICANT: KINZLER, KENNETH
APPLICANT: MARKHAM, ALEXANDER F.
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: THILVERIS, ANDREW
TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
NUMBER OF SEQUENCES: 102
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allegretti, LTD
STREET: 1001 G Street, NW
CITY: Washington

```

STATE: D.C.
COUNTRY: USA
ZIP: 20001-4598
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/289,548A
FILING DATE: 12-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 1107,46943
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2842 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: APC
US-08-289-548A-7

```

Query Match 1.6%; Score 120.5; DB 1; Length 2842;

Best Local Similarity 19.8%; Pred. No. 0.46; Matches 184; Conservative 104; Mismatches 328; Indels 311; Gaps 47;

```

QY 34 GOLMPETLKSPFPCNPKLVNMPKSVLSKPSDIKFAAIGNLEIPDPGTDLEK-92
DB 1920 GQPKPIQKSTFP-----QSSDI-----PDRGAATBEKL 1950
QY 93 QDWTERPQVCM---GVMTVLSDTIR-----YFSPSVPM-----123
DB 1951 QNFAIENTPVCFSHNSLSLSDIDQNNKNEPIKETEPDQSGSPKQASGVAPKS 2010
QY 124 -----PVC---HTGKQVIPHDAEDIMQAEIV-----RNM 152
DB 2011 FHVEDTPVPCSRSSLSLSIDSEDL---LQECISSAMPKKKKPSRLKGDNEKSRNM 2067
QY 153 -----KENLQADPQ-----FDWRLI-----NVFSPNASQ---CYLCPESAQ 185
DB 2068 GGLIGEDLTLDKDIQRPDEHGLSPSENFDMKALQEGANSIVSLHQAAAAACLSRQ- 2126
QY 186 NGLAAGVDLMGVLDYLOQEVPRPAFVNLVLSVAEVSROYHGTWLSPAPEPCNSEET 245
DB 2127 ---ASSDSDBILSLKSGISLSP---FHLTPDQEKPKPTSNK-----GPRILKPEKKS 2173
QY 246 TRLAKVVMQSY-----QEAMNSILASRSRSEQSFVTFQPPFYETTPSLHSEDPQLD 300
DB 2174 TLETKEIIESSKIKGKQKVKSLITGKRVANSSEISQOMQOP-----LQANMPSISR 2225
QY 301 STTLAMHLMRMMEPAGEKEPPLSVKRGKRPMPKCPQSPFLFSRYRNSNYLTRLOKPDKL 350
DB 2226 GRIMI-HIRG--VRNSSSSSTSPVS-KKGPKLKTPAKSP--SEGQRTTSPRGAKS VKS 2279
QY 361 EV---REGAIRCPCDPKS-----DTVPTSVHRLKPADINVIQALGDSITAGNGAGSTP 411
DB 2280 ELSPVARSQTSIQGSSKAPRSRSGRSDSTSPRPAQ-QPLSRPIQSPGNNSISPGNGISPP 2338
QY 412 GNVLDVLTVRGISWSVGDEENIGTVTTLANILREFPNSIKGF-SVETGR--ETSPNAFL 468
DB 2339 NK-----LSQLPRTSSPTASTYKSSGSGKMSYTSPPGRM 2372
QY 469 NQAVAGRAEDLPVQARLVDLMKNDTRHFOED-WKITLFLFGNDLQFCFNDLVHYSP 527

```

```

DB 2373 SQ-----QNLTKQ-----TGLSKMASSIPRSSASKGLQNMNGNG-----AAKKVELSR 2417
QY 528 QNFTDNIQKALDLHAEVPRPAFVNLVTVLEIVNLRELQYQKYYCPMILRLSLCPVLKF 587
DB 2418 MSSTKSSGSESD--RSRRP-----LVVRQSTRIKE-----APSPTR-----2452
QY 588 DDMSTELATLIEPNKKFQEKTHQLIESGRDYDTEDFVTVVQPFENVDMKTEGSLPDNS 647
DB 2453 -----RKLEBSASFESLSPSSRPASPTRSQAQTVLSPSLPD--2489
QY 648 FPADECFHSSKSHSPASALMNM---LEPV-----GQKTRHKENKINITCPNQVOP 699
DB 2490 -----MSLSTHSSVQAGGMRKLPNNLSPTIEYNDGRAXKHDLA-RSHSSPSRL-P 2539
QY 700 FLATYKSMQGHGTWLPICRDRAPSAHPHSVHALRPADIVVAALGDSLTAAGISGKPD 759
DB 2540 INRS-----GTWGREHSHSSSLPRVSTWRTSSSSILSASSSESEK-----AKSE 2586
QY 760 DLPDVTTVYRGSLSYSGDGSLENTVTLPIILR-----EFN-RNLGYAVGTGDANDTNA 813
DB 2587 DEKHVNS-----ISGTQSKKENQVSAKGTWRKIKENEFSPNISTSQTVSSGATN-----2635
QY 814 FLNQAVGAKAEDLMGQVOTLMQKMD 840
DB 2636 -----GASKTLIYQMAPAVKTED 2655

```

RESULT 10
US-08-452-654-7
Sequence 7, Application US/08452654
Patent No. 5691454

GENERAL INFORMATION:

```

APPLICANT: ALBERTSEN, HANS
APPLICANT: ANAND, RAKESH
APPLICANT: CARLSON, MARY
APPLICANT: GRODEN, JOANNA
APPLICANT: HEDGE, PHILIP J.
APPLICANT: JOSLYN, GEORF
APPLICANT: KINZLER, KENNETH
APPLICANT: MARKHAM, ALEXANDER F.
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: THILVERIS, ANDREW
TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESSES:
ADDRESS: Banner, Birch, McKie & Beckett
STREET: 1001 G Street, NW
City: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001-4598
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/452,654
FILING DATE: 25-MAY-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/741,940
FILING DATE: 08-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 1107,035574
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
INFORMATION FOR SEQ ID NO: 7:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 2842 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: APC
US-08-452-654-7

```

```

Query Match 1.6%; Score 120.5; DB 1; Length 2842;
Best local Similarity 19.8%; Pred. No. 0.46;
Matches 184; Conservative 104; Mismatches 328; Indels 311; Gaps 47;

```

```

QY 34 GQLMPELTNKPCCPNKLGVMNPKSVHSLKPSDIKFAVAIGMLEIPDPGTGLEK-92
DB 1920 GQKPFILQKOSTFP-----QSSKDI-----PDRGAATDEKL 1950
QY 93 QDWTERPQOCM---GVMTVLSDIIR-----YFSPSVPM-----123
DB 1951 QNPAIENTPVCFSHNSLSLSDIDENNKNENEPKETEPDPDSQGEPSKPAQSVAPKS 2010
QY 124 -----PVC---HTGKRVIPHDGAEDLWIAQELV-----RNM 152
DB 2011 FHVEDPVCFSRNSLSLSLSDIDEDDL---LQECISAMPKKKKPSRLKGDNEKHSRRM 2067
QY 153 ---KENLQDLPQ-----PDWKL1---NVFSPNASQ---CYLCPSAQ 185
DB 2068 GGILGEDLTLDLKDIOEPDEHGLSPDENFDMKALQEGANSIVSSIHQAAMAAACLSRQ- 2126
QY 186 NGLAGGVEELMGVLDVLAQEVPRAPVNIYDLSEVAVSQHYGTMLSPAPCECNCSEET 245
DB 2127 ---ASDSDBLSLKGISLISP---FHLTPDQEKPEPTSNK-----GPRILKDEKS 2173
QY 246 TRLAKVVMQWSY-----QEAANSLASRYSEQSFVVFPQPFYETTPSLHSEDPRLQD 300
DB 2174 TLETKKIESSEKIKGGKYYKSLITGKAVNSEISQGMQP-----LQANMPSISR 2225
QY 301 STTLAHLNMRMEPAGEKDEPLSVKRGKPMKCPQSEPYLFSYRNSNYLTRLOKQDKL 360
DB 2226 GRIMI-HIIP--VRNSSSTSPVS--KKGPRPKTPASKP--SEGQVATSPRGAKEVKS 2279
QY 361 EV---REGAEIKCPDQPS-----DTPVSVHRLKPADINIVGALCDSITAGACSTP 411
DB 2280 ELSPVARQTSQIGSSVAPSRSGSRSTSPRPAQ-QPLSRPIQSPGRNISTSPRNGISPP 2338
QY 412 GNVLDVLTQYRGLSWSVGDENIGVTTLANILREFNPSLKGF-SVGTGK--ETSPNAFL 468
DB 2339 NK-----LSQLRTPSSPSTASTKSSGSGMSTSPERGOM 2372
QY 469 NQAVAGRAEDLPVQARLVLDLKNDRIRIHOED-WKITLFLTGNDLCPGNDLVHYS 527
DB 2373 SQ-----QNLTKQ-----TGLSKNASSIPRSASAKGLQNMNGNG---AAKVELSR 2417
QY 528 QNTDNIIGKADLIHAVEPRAPVNIYDLSEVAVSQHYGTMLSPAPCECNCSEET 567
DB 2418 MSSTKSSGSESD--RSERP-----VLVROSTRIKE-----ASPFLIR 2452
QY 588 DNDSTELATLIEFNKKFOEKTQHOLIESGRVDTREDFTVVVQPFENVDMKTESEGLPDNS 647
DB 2453 -----RKLIESASFESLSPSRPASPTRSQAQTPVLSPLPD--2489
QY 648 FPAIDCFHSSKSHSRAASALMNM---LEPV-----GQKTRHKEKNKINITCPNOVOP 699
DB 2490 -----NSLSTHSSVQAGWKRKLPPNLSPTIEYNDGRPAKHDA-RSHSSPSRL-P 2539
QY 700 FLATYKNSQGHCTWLPBCRRAPSAHPTSVHALRPADIVVAAGDSLTAAGCIGSKPD 759
DB 2540 INNS-----GTWKRREHSHSSSLPRVSTWRTTGGSSILASSSSEK-----AKSE 2586
QY 760 DLDPVTTQYRGLSWSVGDENIGVTTLANILREFNPSLKGF-SVGTGK--ETSPNAFL 813

```

```

DB 2587 DEKIVNS-----ISGTRQSKENQVSAKGTWRIKINENFSSPTNSTSQTVSSGATN-----2635
QY 814 FLNQAVPQAKAEDLMSQVOTLMQXMD 840
DB 2636 -----GASKTLIYQMAPAVAKTED 2655

```

RESULT 11

```

US-08-449-731-7
Sequence 7, Application US/08449731
Patent No. 6413727
GENERAL INFORMATION:
APPLICANT: ALBERTSEN, HANS
ANAND, RAKESH
CARLSON, MARY
GRODEN, JOANNA
HEDGE, PHILIP J.
JOSLYN, GEOFF
KINZLER, KENNETH
MARKHAM, ALEXANDER F.
MAKAMURA, YUSUKE
THALIVERTIS, ANDREW

```

```

TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
GENE IN COLORECTAL CANCER IN HUMANS

```

```

NUMBER OF SEQUENCES: 102
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner & Allegretti, LTD
STREET: 1001 G Street, NW
CITY: Washington
STATE: D. C.
COUNTRY: USA
ZIP: 20001-4598

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,731
FILING DATE: 25-May-1995
CLASSIFICATION: <Unknown>

```

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/289,548
FILING DATE: 12-AUG-1994

```

```

ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 1107.46943

```

```

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299

```

```

INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 2842 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
IMMEDIATE SOURCE:
CLONE: APC
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-08-449-731-7

```

```

Query Match 1.6%; Score 120.5; DB 4; Length 2842;
Best local Similarity 19.8%; Pred. No. 0.46;
Matches 184; Conservative 104; Mismatches 328; Indels 311; Gaps 47;

```

```

QY 34 GQLMPELTNKPCCPNKLGVMNPKSVHSLKPSDIKFAVAIGMLEIPDPGTGLEK-92
DB 1920 GQKPFILQKOSTFP-----QSSKDI-----PDRGAATDEKL 1950

```

```

QY 93 QDWTERRQOYCM---GVMTVLSDIIR-----YFSPVPM----- 123
D 1951 QNPAIENTPVCFSHNSLSLSDIDQENNNKENEPKETEPDPSQSPKQASGYAPKS 2010
QY 124 -----PVC---HTGKRVIPHOGAEDLMIOAQLV----- 152
D 2011 FHVEDTFCFSRNSLSLSLSDIDEDL---LQECISSAMPKPKKPRSLKGDNEKHSPPRM 2067
QY 153 -----KENVLODFQ-----FDKLI-----NVFSSNASQ---CYLCPSAQ 185
D 2068 GGLGEDLTLDLKDIOQPDSEHGLSPDSENFDMWKAQEGANSIVSSLHQAACLSRQ- 2126
QY 186 NGLAAGVDELMDGLDYLOQEVPRAFVNLVDLSEVAEVSROYGTWLSPAPEPCNSEET 245
D 2127 ---ASSDSDBLSLKSIGSLGSP---FHLTPDQEKPTSNK-----GPRILKPEKS 2173
QY 246 TRLAKVVMQMSY-----QEAANSILASSRYSBOESTVVPQPFYETTPSLHSEDRLOD 300
D 2174 TLETKEIESBSKGIKGGKVKYKSLITGKVRNSNEISGQMKQ-----LQANMPSISR 2225
QY 301 STLAWHLNRMMEPAGEKDEPLSVKGRPMKCPQSQSPYLFYSRNSNYLTRLOKPODKL 360
D 2226 GRTMI-HIPG--VRNSSSTSPVS-KKGPKLTPAKSP--SEGQATTSPRGAKSPVKS 2279
QY 361 EV-----REGAIRCPODPS-----DTVPTSVMRLKPADINVIGALGDSLTAAGAGSTP 411
D 2280 ELSFVARQTSQIGSSKAPBSRSGSDSTPSRPAQ-QPLSRPIQSPGRNNSISPRGNGISPP 2338
QY 412 GNVLDVLTQYRGLSWSVGDENIGTVTTLNLIREFNSLKGK-SVGTGK--ETSPNAFL 468
D 2339 NK-----LSQLPRTSPSTASTSSSGSKMSTYSPGRQ 2372
QY 469 NOAVAGRADLPLVQARLVDMKNDTRHFOED-WKIIITLFIGNDLCPDNDVHYS 527
D 2373 SQ-----ONLTQKQ-----TGLSKNASIPRSESASGLQNMNGNG---ANKKVELSR 2417
QY 528 QNFDTNIGKALDILHAEPRAFAVNLVLEIVNRELVOEKVYCPMILRSLCPVLKLF 587
D 2418 MSSTKSSGSESD-RSERP-----VVRQSTFIKE-----APSPILR----- 2452
QY 588 DDNSTELATLIEFNKKFOEKTHOLIESGRYDTREDFTVVOPEFENVDMKTESGLPDS 647
D 2453 -----RKLSESASFESLSPSRPASPTPSQAQTVLSLSPD-- 2489
QY 648 FPAPDCHESSKSHRASAALNMNM---LEPV-----GQKTRHKEFNKINITCPNOVOP 699
D 2490 -----MSLSTHSSVQAGWRLPNNLSPTIEYNDGRPAKRHDIA-RSHSESPRL-P 2539
QY 700 FLRTYKNSMOGHGWLPCRDAPALHPTSVHALRPADIOVVALGDSLTAAGNGISGKPD 759
D 2540 INNS-----GTWKEHSHKSSSLPRVSTWRTSSSSSILSSASSSEK-----AKSE 2586
QY 760 DLDPTVTQYRGLSYSAGDGLSENVTLIPILR-----EFN-RNLGYAVGTGDANDTNA 813
D 2587 DEKRVNS-----ISGTQSKENQVSAKGTWRIKENEFSPTNSTGQTVSSGATN----- 2635
QY 814 FLNOAVPGAAEDIMSOVQTLMOKMD 840
D 2636 -----GAESKTLIYQMAVAVSKTED 2655

```

RESULT 12
US-07-741-940-2
Sequence 2, Application US/07741940
Patent No. 5352775

GENERAL INFORMATION:
APPLICANT: ALBERTSEN, HANS
APPLICANT: ANAND, RAKESH
APPLICANT: CARLSON, MARY
APPLICANT: GRODEN, JOANNA
APPLICANT: HEDGE, PHILIP J.
APPLICANT: JOSLYN, GEROFF

```

? APPLICANT: KINZLER, KENNETH
? APPLICANT: MARKHAM, ALEXANDER F.
? APPLICANT: NAKAMURA, YUSUKE
? APPLICANT: THIVERIS, ANDREW
? TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
? TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
? NUMBER OF SEQUENCES: 94
? CORRESPONDENCE ADDRESS:
? ADDRESS: Banner, Birch, McKie & Beckett
? STREET: 1001 G Street, NW
? CITY: Washington
? STATE: D.C.
? COUNTRY: USA
? ZIP: 20001-4598
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/07/741,940
? FILING DATE: 19920109
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Kagan, Sarah A.
? REGISTRATION NUMBER: 32,141
? REFERENCE/DOCKET NUMBER: 1107.035574
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 202-508-9100
? TELEFAX: 202-508-9299
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2843 amino acids
? TYPE: AMINO ACID
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-07-741-940-2

```

Query Match 1.6%; Score 120.5; DB 1; Length 2843;
Best Local Similarity 19.8%; Pred. No. 0.46;
Matches 184; Conservative 104; Mismatches 328; Indels 311; Gaps 47;

```

QY 34 GQLMPELTLSNPPCPNKLGVNMPKSVSLKPSDIKFAVLAIGNLEIPDPGTDLEK- 92
D 1921 GQPKPLQKSTFP-----QSSKDI-----PDRGAATDEKL 1951
QY 93 QDWTERRQOYCM---GVMTVLSDIIR-----YFSPVPM----- 123
D 1952 QNPAIENTPVCFSHNSLSLSDIDQENNNKENEPKETEPDPSQSPKQASGYAPKS 2011
QY 124 -----PVC---HTGKRVIPHOGAEDLMIOAQLV----- 152
D 2011 FHVEDTFCFSRNSLSLSLSDIDEDL---LQECISSAMPKPKKPRSLKGDNEKHSPPRM 2067
QY 153 -----KENVLODFQ-----FDKLI-----NVFSSNASQ---CYLCPSAQ 185
D 2068 GGLGEDLTLDLKDIOQPDSEHGLSPDSENFDMWKAQEGANSIVSSLHQAACLSRQ- 2126
QY 186 NGLAAGVDELMDGLDYLOQEVPRAFVNLVDLSEVAEVSROYGTWLSPAPEPCNSEET 245
D 2127 ---ASSDSDBLSLKSIGSLGSP---FHLTPDQEKPTSNK-----GPRILKPEKS 2173
QY 246 TRLAKVVMQMSY-----QEAANSILASSRYSBOESTVVPQPFYETTPSLHSEDRLOD 300
D 2175 TLETKEIESBSKGIKGGKVKYKSLITGKVRNSNEISGQMKQ-----LQANMPSISR 2226
QY 301 STLAWHLNRMMEPAGEKDEPLSVKGRPMKCPQSQSPYLFYSRNSNYLTRLOKPODKL 360
D 2227 GRTMI-HIPG--VRNSSSTSPVS-KKGPKLTPAKSP--SEGQATTSPRGAKSPVKS 2280
QY 361 EV-----REGAIRCPODPS-----DTVPTSVMRLKPADINVIGALGDSLTAAGAGSTP 411
D 2281 ELSFVARQTSQIGSSKAPBSRSGSDSTPSRPAQ-QPLSRPIQSPGRNNSISPRGNGISPP 2339

```

```

QY 412 GNVLDVLTQYRGLSWSVGDENIGVTTLANTLREFPNLSKGF-SVGTGK--ETSPNAFL 468
Db 2340 NK-----LSQLPRTSSPSTASTKSSGSGMSTSTSGRQM 2373
QY 469 NQAVAGRAEDLPVQARRLVLDLMDKDRTHFOED-WKITLFTGNDLDFCNDLVHYSP 527
Db 2374 SQ-----QNLTKQ-----TGLSKNASSIPRSESASGKLQNMNNGNG---ANKKYELSR 2418
QY 528 QNFTDNIGKALDILHAEVPRAFVNLVTVLEIVNLRELQYQKYYCPRMLIRSLCPVLKP 587
Db 2419 MSSTKSSGSESD--RSERP-----VLVRQSTFIKE-----APSPILR----- 2453
QY 588 DNSTELATLIEFNKKFOEKTQHLIESGRYDTRDFVTVVQPFENVDMPTSEGLPDNS 647
Db 2454 -----RKLIESASPESSPSRSPASPTRSQAQTPVLSPLPD-- 2490
QY 648 FPAIDCFHSSKSHSRAASALMNNM--LEPV-----GQKTRHKEENKINITCPNOVOP 699
Db 2491 -----MSLSTHSSVQAGWRKLPNLSPTIEYNDGRPAKHDLA-RSHSESPRL-P 2540
QY 700 FLRTYKNSQMGHGTWLPCCRBAASALHPTSVHALRPADIQVVAALGDSLTAAGIGSKPD 759
Db 2541 INNS-----GTWKRHSKSHSSLPVSTWRTGSSSILSASSESEK-----AKSE 2587
QY 760 DLPDVTTOYRGLSYSAGDGLSNVTTLPNILR-----EFN-RNLGVAVGTGDANDTNA 813
Db 2588 DEKHVNS-----ISGTQSKENQVSAKGTWKIKENEFSPNISTSQVSSGATN----- 2636
QY 814 FLNQAVPQAKAEDLMSQVOTLMQMKD 840
Db 2637 -----GAKSKTLIYQMAPAVSKTED 2656

```

```

RESULT 13
US-08-289-548A-2
; Sequence 2, Application US/08289548A
; Patent No. 5648212
; GENERAL INFORMATION:
; APPLICANT: ALBERTSEN, HANS
; APPLICANT: ANAND, RAKESH
; APPLICANT: CARLSON, MARY
; APPLICANT: GRODEN, JOANNA
; APPLICANT: HEDGE, PHILIP J.
; APPLICANT: JOSLYN, GEOFF
; APPLICANT: KINZLER, KENNETH
; APPLICANT: MARKHAM, ALEXANDER F.
; APPLICANT: NAKAMURA, YUSUKE
; APPLICANT: THLIVERIS, ANDREW
; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
; NUMBER OF SEQUENCES: 102
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti, LTD
; STREET: 1001 G Street, NW
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001-4598
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/289,548A
; FILING DATE: 12-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kagan, Sarah A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 1107,46943
; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2843 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-289-548A-2

Query Match 1.6%; Score 120.5; DB 1; Length 2843;
Best Local Similarity 19.8%; Pred. No. 0.46;
Matches 184; Conservative 104; Mismatch 328; Indels 311; Gaps 47;

QY 34 GQMPETLKNPPCPNKLGVNMPKSHSLKPSDIKVAALGNLEIPDPDGLK- 92
Db 1921 GQPKPLQKOSTP-----QSSKDI-----PDRGAATDEKL 1951
QY 93 QDWTERPQOQCM---GVMTVLSDIIR-----YFSPSVPM----- 123
Db 1952 QNPALIENTVPCFSHNSLSLSDIDQENNNKNEPIKTEPPDQSEPKPQASGAPKS 2011
QY 124 -----PVC---HTGKRVIPIHGAEDLMTQAOELV-----RNM 152
Db 2012 FIVEDTPVCFSRNSLSLSDIEDL--LQECISSAMPKKKPERLKDNKHSPPRM 2068
QY 153 -----KENLDLPQ-----FMKLI-----NVEFSASQ--CYLCSAQ 185
Db 2069 GGLIGEDLTLTDLKDQRPDSEHGLSPSENFMKALQEGANSTVSLHQAALAAACLSRQ- 2127
QY 186 NGLAAGVDELGVLDYLQOEVPRAFNVLSEVAEVSROYHTWLSAPAPCNCSEET 245
Db 2128 ---ASSDSDSLSLKSGISLGSF--PHLTPDQEKFTSNK-----CPRLKRGKS 2174
QY 246 TRLAKVMQMSY---QEAWSLASSRYSQESFTVVPQPFYETTSLSHSEDRLOD 300
Db 2175 TLETKKIESESKIKGKKVYKSLTGKVRNSNIEIGQKKQ-----LQNMPSISR 2226
QY 301 STTLNHLNRMMEPGEDEPLSVHGRPMKCPSESPLYBSRYNSNLTIRQKQDYL 360
Db 2227 GRMTI-HIPG--VRNSSSSTSPVS--KKGPPLTPPARKSP--SEGQATTSPPRAKSVKS 2280
QY 361 EV---REGAEIRCPDKPS-----DVTPTSVAHLRPADINVAGLDSLTAAGAGSTP 411
Db 2281 ELSPVAKQTSQIGGSKABSRGSRDSTPSRAQ-QPLSRPIOSPARNISPEARNISPE 2339
QY 412 GNVLDVLTQYRGLSWSVGDENIGVTTLANTLREFPNLSKGF-SVGTGK--ETSPNAFL 468
Db 2340 NK-----LSQLPRTSSPSTASTKSSGSGMSTSTSGRQM 2373
QY 469 NQAVAGRAEDLPVQARRLVLDLMDKDRTHFOED-WKITLFTGNDLDFCNDLVHYSP 527
Db 2374 SQ-----QNLTKQ-----TGLSKNASSIPRSESASGKLQNMNNGNG---ANKKYELSR 2418
QY 528 QNFTDNIGKALDILHAEVPRAFVNLVTVLEIVNLRELQYQKYYCPRMLIRSLCPVLKP 587
Db 2419 MSSTKSSGSESD--RSERP-----VLVRQSTFIKE-----APSPILR----- 2453
QY 588 DNSTELATLIEFNKKFOEKTQHLIESGRYDTRDFVTVVQPFENVDMPTSEGLPDNS 647
Db 2454 -----RKLIESASPESSPSRSPASPTRSQAQTPVLSPLPD-- 2490
QY 648 FPAIDCFHSSKSHSRAASALMNNM--LEPV-----GQKTRHKEENKINITCPNOVOP 699
Db 2491 -----MSLSTHSSVQAGWRKLPNLSPTIEYNDGRPAKHDLA-RSHSESPRL-P 2540
QY 700 FLRTYKNSQMGHGTWLPCCRBAASALHPTSVHALRPADIQVVAALGDSLTAAGIGSKPD 759
Db 2541 INNS-----GTWKRHSKSHSSLPVSTWRTGSSSILSASSESEK-----AKSE 2587
QY 760 DLPDVTTOYRGLSYSAGDGLSNVTTLPNILR-----EFN-RNLGVAVGTGDANDTNA 813
Db 2588 DEKHVNS-----ISGTQSKENQVSAKGTWKIKENEFSPNISTSQVSSGATN----- 2636

```


QY 814 FLNQAVGAKAEDLMSQVOTLMQMKD 840
 Db 2637 -----GAESKTLIYQMAPAVSKTED 2656

RESULT 14
 US-08-452-654-2
 ; Sequence 2, Application US/08452654
 ; Patent No. 5691454
 ; GENERAL INFORMATION:
 ; APPLICANT: ALBERTSEN, HANS
 ; APPLICANT: ANAND, RAKESH
 ; APPLICANT: CARLSON, MARY
 ; APPLICANT: GRODEN, JOANNA
 ; APPLICANT: HEDGE, PHILIP J.
 ; APPLICANT: JOSLYN, GEOFF
 ; APPLICANT: KINZLER, KENNETH
 ; APPLICANT: MARKHAM, ALEXANDER F.
 ; APPLICANT: NAKAMURA, YUSUKE
 ; APPLICANT: THIVERIS, ANDREW
 ; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
 ; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
 ; NUMBER OF SEQUENCES: 94
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Banner, Birch, McKie & Beckett
 ; STREET: 1001 G Street, NW
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20001-4598
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/452,654
 ; FILING DATE: 25-MAY-1995
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/741,940
 ; FILING DATE: 08-AUG-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kagan, Sarah A.
 ; REGISTRATION NUMBER: 32,141
 ; REFERENCE/DOCKET NUMBER: 1107.035574
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-508-9100
 ; TELEFAX: 202-508-9299
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2843 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-452-654-2

Query Match 1.6%; Score 120.5; DB 1; Length 2843;
 Best Local Similarity 19.8%; Pred. No. 0.46;
 Matches 184; Conservative 104; Mismatches 328; Indels 311; Gaps 47;

QY 34 GQMLPELTKNSPPPCNKNKGAMPSKVSLSKPSDIKPAVAIGNLEIPDPGIGDEK-92
 Db 1921 GQKPKPILOKQSTFP-----QSSKDI-----PDRGATDEKL 1951

QY 93 QDWTERRQVCM---GVMTVLSDIIR-----YFSPVPM-----123
 Db 1952 QNPAIENTPVCFSHNSLSLSDIDQNNKNENPIKETEPDPSQGFSPKQASGVAPKS 2011

QY 124 -----PVC---HTGKRVIPHDGAEGLMIAQELV-----RNM 152
 Db 2012 FHVEDFTVCGRNSLSLSLSDSDL---LQCTISAMPKPKKPRKLGDNKHSPRNM 2068

QY 153 -----KENLODPQ-----FDWKL-----NVEFSNAQ--CYLCPSAQ 185
 Db 2069 GGILGEDLTLDLKDIOIPDSEHGLSPDSENFDMKALQEGANSIVSSLHQAAAAACLSRQ- 2127

QY 186 NGLAAGVDLMGVLDYLOQEVPAFVNLVLSVAEVSQYHGTWLSPAPECNCSSEET 245
 Db 2128 -----ASSDSDSILSLKGSILGSP--FHLFPDQEKKEPFTSNK-----GPRILKGEKS 2174

QY 246 TRLAKYVMQSY-----QENANSLASRYSSEDSFTVVFQPFYETTPSHSEDRLOD 300
 Db 2175 TLRTKIESEKIGKGGKRYKSLITGKVRNSNLSIGOMKQ-----LQANMPSISR 2226

QY 301 STLAWHLNMRMPEAGEKDEPLSVKHGRPMKCPGSEPLTFGRSNVYLRQCPQDKL 360
 Db 2227 GRMTI-HIRG--VRNSSSTSPVS-KKGPKLTPAKSP--SGQRTATSPRAKPSVKS 2280

QY 361 EV-----REGAETRCPPDKPS-----DTVPTSVAHLKPADINVTIGALGDSLTAAGAGSTP 411
 Db 2281 ELSPVARQTSQIGSSKAPBSGSRDSTPSRPAQ-QPLSRPIQSPGRNLSISGRNGISPP 2339

QY 412 GNVLDVLTQYRGLSNVSGDENIGYTTLANILRENPISLKG--SVGTGK--ETSRAFL 468
 Db 2340 NK-----LSQLEPRTSSPSTASTKSSGSGKMSYSPROM 2373

QY 469 NOAVAGRAEDLPVQARLVLDLMDKNDTRIHFQED-WKITLFTIGNDLDFCNDLVHYS 527
 Db 2374 SQ-----ONLTKQ-----TGLSKNASSIPRSASAKGLMOMNNGG---ANKVYELSR 2418

QY 528 QNFTDNGKALDILAEVPAFVNLVLEIYNRELQOEKVVCPRMILRLSLCPVLKP 567
 Db 2419 MSSTKSSGSESD--RSERP-----VLVRQSTRIKE-----APSPFLR 2453

QY 588 DNDSTELATLIEPNKQFOEKTQHLISSGRYDTRREDTTVVQPFPEVNDPKTSEGLPDNS 647
 Db 2454 -----RKLEBSASFESLSPSSRPASPTPSQAQTIVLSPSLPD-- 2490

QY 648 FPAEDCFHSSKSHSRAASALMNM---LEPV-----GOKTRHKKENKINICPNQVOP 699
 Db 2491 -----MSLSTHSSVQAGCRKLPPNLSFTIEYDGRPAKHDA-RSHSSPSRL-P 2540

QY 700 FLRTYKNSMQHGTWLPCCRDAFSALHPTSVAHLRADIQVVAALGDSLTAAGIGSKPD 759
 Db 2541 INRS-----GTWKRHSKHSSSLPRVSTWRTGSSSSILSASESSBK-----AKSE 2587

QY 760 DLPDVTYTRGLSYSGDGSLENVTLPLNLR-----EPN-RNLGYAVGTADADTNA 813
 Db 2588 DEKHVNS-----ISGTQSKENOVSAKGTWRKIKENERSPTNSTSQTVSSGATN----- 2636

QY 814 FLNQAVGAKAEDLMSQVOTLMQMKD 840
 Db 2637 -----GAESKTLIYQMAPAVSKTED 2656

RESULT 15
 US-08-452-655B-2
 ; Sequence 2, Application US/08452655B
 ; Patent No. 5783666
 ; GENERAL INFORMATION:
 ; APPLICANT: ALBERTSEN, HANS
 ; APPLICANT: ANAND, RAKESH
 ; APPLICANT: CARLSON, MARY
 ; APPLICANT: GRODEN, JOANNA
 ; APPLICANT: HEDGE, PHILIP J.
 ; APPLICANT: JOSLYN, GEOFF
 ; APPLICANT: KINZLER, KENNETH
 ; APPLICANT: MARKHAM, ALEXANDER F.
 ; APPLICANT: NAKAMURA, YUSUKE
 ; APPLICANT: THIVERIS, ANDREW
 ; TITLE OF INVENTION: INHERITED AND SOMATIC MUTATIONS OF APC
 ; TITLE OF INVENTION: GENE IN COLORECTAL CANCER IN HUMANS
 ; NUMBER OF SEQUENCES: 102
 ; CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner & Witcoff, Ltd.
 STREET: 1001 G Street, NW
 CITY: Washington
 STATE: D. C.
 COUNTRY: USA
 ZIP: 20001-4598
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/452,655B
 FILING DATE: 25-MAY-1995
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/289,548
 FILING DATE: 12-AUG-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/741,940
 FILING DATE: 08-AUG-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Kagan, Sarah A.
 REGISTRATION NUMBER: 32,141
 REFERENCE/DOCKET NUMBER: 1107,49964
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-508-9100
 TELEFAX: 202-508-9299
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2843 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-452-655B-2

Query Match 1.6%; Score 120.5; DB 1; Length 2843;

Best local Similarity 19.8%; Pred. No. 0.46;

Matches 184; Conservative 104; Mismatches 328; Indels 311; Gaps 47;

34 GQMPETLKNSPPCNPNKLGVMMPKSVHSLKPSDIKVAIAGNLEIPDPGTDLEK-92
 1921 GQPKPLIQKQSTPP-----QSSKDI-----PDGAAATDEKL 1951
 93 QDWTBRQOYCM---GVMTLSDILR-----YFSPSVPM-----123
 1952 QNPAIENTPVCFSHNSLSLSDIDENNKNENPIKETEPDPSQGEPSKPOASGYAPKS 2011
 124 -----PVC---HTGKRVIPHDGAEDLWIOAELV-----RNM 152
 2012 FHVEDTPVCSRSRNSLSLSIDSEDDL---LQECISAMPKKKKPSRLKGDNEKHSFRNM 2068
 153 ---KENLQIDFQ-----FDWKLI---NVFFSNASQ---CYLCPSAQ 185
 2069 GGIIGEDLTLDLKDIOIPDSEHGLSPDSENFWDKAIQEGANSIVSLHQAAAAACLSRQ- 2127
 186 NGLAAGVDELMGVLDYLOQEVPRAPFNLYVDSLVAEVSQYHGTWLSPPAPPCNCSEET 245
 2128 ---ASSDSDSLKSLKGISLGSF---FHLTPDOEKPTSNK-----GPRILKPEKS 2174
 246 TRLAKVVMQWSY-----QAMNSLLASSRYSEQESFTVVFOFPFYEETTPSLHSEDPRLQD 300
 2175 TLETKIESSEKIKGKAKYKSLITCKVNSNEISGMKQP-----LOANMPSISR 2226
 301 STTLAHLNRMMEPAGEKDEPLSVKRGKPMKCPQSEPYLFSYRNSNYLTRLOKPODKL 360
 2227 GRIMI-HIPG--VRNSSSSTSPVS-KKGPLKTPASKSP--SEGQFATTSPPGAKPSVKS 2280
 361 EV-----REGAEIKCPDQDS-----DTVPTSVRLKPADINVGALGDSLTAAGAGSTP 411
 2281 ELSPVAKQTSIQIGSSAPRSRSGSRDSTSPRAPO-QLSRPIQSPGRNNSISGRNGISPP 2339
 412 GNVLDVLTQYRGLSWSVGSDENIGVTTLANIREFNPSLKGK-SVOTGK--ETSPNAFL 468

2340 NK-----LSQLRTSSPSASTKSSGSGMGSYTSPGRQM 2373
 469 NQAVAGRAEDLPVQARLVLDLMKNDTRJHFQED-WKITLFLGNDLDFCNDLVHYSP 527
 2374 SQ-----QNLTKQ---TGSKNASLIPRSASAKGLQNNNGNG---ANKVELSR 2418
 528 QNFTDNIGKALDILHAEPRAFNLTVEIIVLRELIOEKKYCPRMILRSICPVILKF 587
 2419 MSTKSSGSESD--RBERP-----VLVROSTIKK-----ABSPTLR-----2453
 588 DDNSTELATLIEFNKKFOEKTQILIESGRYDREDFTVVQPFENVDMPTSEGLPDNS 647
 2454 -----RKLIESASFESLSPSSRPASPTRSQAQTFVLSPLPD--2490
 648 FFAPDCEHSSKSHSRAASALMNM---LEPV-----GQKTHKFNKNITCPCNOVOP 699
 2491 -----MSLSTHSSVQAGMKRLPPNLSPTIEYNDGRPAKHDA-RSHSSEPSRL-P 2540
 700 FLRTYKNSMQGHGTWLPCCRDRAPSLHPTSVAHLPADIQVVAALGDSLTAAGIGSKPD 759
 2541 INRS-----CTWIKREHSHSSSLPRVSTWRTGSSSILSASSESEK-----AKSE 2587
 760 DLPDVTTOYRGLSYSAGDGSLENVTLTPNILR-----EFN-RNLGYAVGTGDANDTNA 813
 2588 DEKHVNS-----ISGTQSKENQVSAGTWRKIKENEFSPNTNSTSQTVSSGATN-----2636
 814 FLNQAVGAKAEDLMSQVOTLMQKMD 840
 2637 -----GAEKTLIYQMAPAVSKTED 2656

Search completed: January 6, 2004, 19:37:43
 Job time : 44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 6, 2004, 18:14:14 ; Search time 100 Seconds
(without alignments)
3762.407 Million cell updates/sec

Title: US-10-054-691-2
Perfect score: 7766
Sequence: 1 MGIRPGIFLLELLLGQCT.....RCRRGGRREDPPMSLRVAL 1458

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: SP archaea:*
 - 2: SP bacteria:*
 - 3: SP fungi:*
 - 4: SP human:*
 - 5: SP_invertebrate:*
 - 6: SP_mammal:*
 - 7: SP_mhc:*
 - 8: SP_organelle:*
 - 9: SP_phage:*
 - 10: SP_plant:*
 - 11: SP_rodent:*
 - 12: SP_virus:*
 - 13: SP_vertebrate:*
 - 14: SP_unclassified:*
 - 15: SP_virus:*
 - 16: SP_bacteriap:*
 - 17: SP_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5624.5	72.4	1463	11	070320
2	5279.5	68.0	1450	11	054728
3	2270	29.2	423	4	081UP7
4	1392	17.9	270	4	096DP9
5	1145	14.7	394	11	09dY6
6	636	8.2	377	5	09N224
7	635	8.2	186	11	08K255
8	628.5	8.1	382	5	023119
9	625.5	8.1	374	5	001300
10	616	7.9	398	5	08MX08
11	584	7.5	424	5	09VM80
12	515	6.6	447	5	09VM08
13	469.5	6.0	460	5	095Q11
14	406.5	5.2	348	5	021799
15	360.5	4.6	528	5	P90862
16	360.5	4.6	565	5	081476

17	359	4.6	484	5	Q21798	Q21798 caenorhabd
18	314	4.0	471	5	Q9N4P5	Q9N4P5 caenorhabd
19	215	2.8	433	5	Q965M0	Q965M0 caenorhabd
20	159.5	2.1	941	16	Q97CZ3	Q97CZ3 clostridium
21	146.5	1.9	1755	16	Q9PC33	Q9PC33 xyella fab
22	145	1.9	2153	12	Q9Y0R5	Q9Y0R5 tula virus
23	141.5	1.8	799	4	Q8N265	Q8N265 homo sapien
24	141.5	1.8	873	5	Q01922	Q01922 caenorhabd
25	139	1.8	9439	16	Q8CP76	Q8CP76 staphylococ
26	137	1.8	4306	11	Q9J179	Q9J179 rattus norv
27	136	1.8	685	16	Q8XST4	Q8XST4 ralsconia s
28	135	1.7	1399	5	Q9V4R0	Q9V4R0 drosophila
29	134.5	1.7	16274	5	Q8IR22	Q8IR22 drosophila
30	133	1.7	1708	17	Q26769	Q26769 mechanobact
31	132	1.7	809	3	Q9HEA5	Q9HEA5 neurospora
32	132	1.7	1181	16	Q9PL06	Q9PL06 chlamydia m
33	131.5	1.7	1003	10	Q94856	Q94856 nicotiana t
34	131	1.7	465	12	Q90729	Q90729 human papil
35	130.5	1.7	891	10	Q9FZ40	Q9FZ40 arabidopsis
36	130.5	1.7	1165	5	Q01923	Q01923 caenorhabd
37	130.5	1.7	1545	5	Q9XX31	Q9XX31 caenorhabd
38	130.5	1.7	2792	5	Q814R2	Q814R2 plasmodium
39	130.5	1.7	2922	5	Q81PJ3	Q81PJ3 drosophila
40	130.5	1.7	3270	4	Q8WWM3	Q8WWM3 homo sapien
41	130	1.7	923	17	Q8U434	Q8U434 pyrococcus
42	129.5	1.7	1245	12	Q9Y709	Q9Y709 atelina her
43	128.5	1.7	974	10	Q8S905	Q8S905 arabidopsi
44	128.5	1.7	1003	10	Q9LP05	Q9LP05 arabidopsi
45	128.5	1.7	2765	11	Q9RIK2	Q9RIK2 rattus norv

ALIGNMENTS

RESULT 1	ID	Query Match	Score	Match	Length	DB ID	Description
070320	070320	72.4%	5624.5	DB 11	Length 1463		
AC	070320						PRELIMINARY; PRT; 1463 AA.
DT	01-AUG-1998 (TEMBLrel. 07, Created)						
DT	01-AUG-1998 (TEMBLrel. 07, Last sequence update)						
DT	01-MAR-2003 (TEMBLrel. 23, Last annotation update)						
DE	Phospholipase B.						
GN	PLB.						
OS	Cavia porcellus (Guinea pig).						
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
OC	Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.						
OX	NCBI_TaxID=10141;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RC	STRAIN=Ctrl.(BPA)BR; TISSUE=Intestine;						
RX	MEDLINE=98256256; PubMed=9593672;						
RA	Delagebeaudet C., Gassama-Diagne A., Nauze M., Ragab A., Li R.Y.,						
RA	Capdevielle J.J., Ferrara P., Fauvel J., Chap H.;						
RT	"Ectopic Epithelial Expression of Guinea Pig Intestinal Phospholipase						
RT	B. Possible role in sperm maturation and activation by limited						
RT	proteolytic digestion";						
RL	J. Biol. Chem. 273:13407-13414(1998).						
DR	EMBL; AF045454; AAC40129.1; -						
DR	InterPro; IPR001087; Lipase_GDSL.						
DR	InterPro; IPR005829; Sug transporter.						
DR	Pfam; PF00657; Lipase_GDSL; 3.						
DR	PROSITE; PS01098; LIPASE_GDSL_SRR; 2.						
DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.						
SQ	SEQUENCE 1463 AA; 162175 MW; 101C4969815F24B1 CRC64;						
QY	1 MGIRPGIFLLELLLGQCTPQIHSTPRKSTLEGQUMPELTAKNSPPCNPNKLGVMPSK 60						
DB	1 MGIRPGIFLLELLLGQCTPQIHSTPRKSTLEGQUMPELTAKNSPPCNPNKLGVMPSK 59						

Qy	61	VSHSLKSPDKEVVAIGNLEIPRDPDGTGEKQDHTER-POOVGWMVTLSDIIRPSP	119
Db	60	SVHTILTPADIKLIAIGMETRPBDSGVAVLDTSEBTEKPRGCMQMTVLSDIISHERP	119
Qy	120	SVMPFVCHTGRKAVIPHDGAEIDLMIQAOBLVRMKENLQDPOFDWKLINFFSNASOCYL	179
Db	120	SVILPPTCPMPSAAVRGVGEELRTOAEELVSLKKNQPLDFOQDWKLINFFSNASLCYL	179
Qy	180	CPSAQONGLAAGVDELNGVLDYLOOEVPRAVNLVDLSEVAEVRQHYGTMUS-PAPER	238
Db	180	CPSAEENPRLMSNMDKLGIILHYHIOEVPRAVNLVDLEEVAMPRMHQTMLSRSPBPA	239
Qy	239	CNCSBETRLAKVVMQMSYOEAAMNSLLASRSSEOSFFVVPQPFYEPTLSHSDPRL	298
Db	240	CGCSEETKLDTVVMQMSYQETWDBLSSSNDDSEFVYVFPQPFYEVSFV--BEPBS	297
Qy	299	QDSTTLAWHLNWRMMEPAGEKDEPISVKGHEPMKCPSOESPYLFSYSRNSNYLTIRLOKPOD	358
Db	298	QDPFTLALSLWNMMKPYGQKDEPPESTIERRMKCPSOESPYLFTYRSNSYGRLLKQOR	357
Qy	359	KLEVRBEGEIRCPDKOPSDTVPTSVHRLKPADINIGALGDSLITANGAGSTPGVNLVYL	418
Db	358	QKHEBEGEIRCPDKOPSDSTPSTSVHRLKPADIKYIGALGDSLITANGAGSGSRPGNILDYL	417
Qy	419	TOYRGLSVSGDEMIQVTTLTANILREBNPBLKQFSYGTGKETSPNPLNOAVVAGBAE	478
Db	418	TEYRGLSVSIGADNHISVTTLPNLILREBNPBLKQFSTGTGANSVGAFENAVVAGRRG	477
Qy	479	DLPOARLAVLDMKNDTRIHFOEDMKIITLPIFGNDLDCDFNDLHYSPONFTDNIIGKAL	538
Db	478	DLIPQARLVLDMKNHTSINFEBDKIITVPIFGNDLDCDFGSDPYTNSEPTNDIMIQAL	537
Qy	539	DILHAEPRAFVNILVTLEIVNLRELIOEKVYICPRMILRSICPVLKFDNSTELATLI	598
Db	538	DILHAEPRAFAFNWVKVQIVNLRELAYKDSRVSCTRLILMRLCRCVLLPDNDSTELASHI	597
Qy	599	EENKRFQOKTHQLESBGRYDPRDETTVVPQPFENWMDPKTSEGLPDMSFAPDCPFHSS	658
Db	598	DINKKYQERTHQLBESGRYDPRDETTVALQPFKEKDIIPKTSSEGLPDNTSFAPDCPFHSS	657
Qy	659	KSHSRLASALNNMMLPEVQOKTTRKPFENKIITCPNQOQPLFRYKMSMOGHGWLPCR	718
Db	658	KTHRAASALMKMMLPEVQOKTTCQNNFENSIDIICPNQAFPLSTLYKNGIBSHGWLPCR	717
Qy	719	DRABSAHLPTSVAHLRPADIOYVAALGDSLITANGSIGSKPDDLPDVTYQYRGLSYSAGSD	778
Db	718	ERTESASPTSVHAIARPADVRVVAALGDSLITAGSIGSKPGLADAVITQYRGLSYSAGSD	777
Qy	779	GLSLNVTTLPNILREBNRLTCYVANGTGDANDTNALNOAVVAGAAQAEIDLMSQVOTLMQK	838
Db	778	GLSLNVTTLPNILREBNRLTGAYAGTGDASNTNPLNOAVVAGAAEELMSQVOTLVQK	837
Qy	839	KODHRVNFHEBMKVITTVILIGSGDLDYCYDSNULYSAANFVHDLRLAALDVLHREVEPVLYN	898
Db	838	KODPRINHEBMKVITTVILIGNDLCNCHTDLDBYSSANFPHNLNALDILHREVEPRALYN	897
Qy	899	LVDLNPITIMQOVELGNBDKCPVOQASVLCNCVTLIRENSOELARLEAFSRAYSRESMREL	958
Db	898	LVDFNPNIPIMQOVELGNBDKCPVOQASLTCNCVLSIRENSVEIARMDALTRAYQSMREL	957
Qy	959	VGSGRYDPTQOEDPSVVLQPPFONIQLPVLADGLPTGSFPAPODCIHNQKFHSQLARALMTN	1018
Db	958	VESGGRYDPRDEPSVVLQPPFLINIRLPILEDGPRDTPSFAPDCINQGFHSQLSALWVN	1017
Qy	1019	MLEPLGSGTTELLDIRAEMPICTPQNEPELTPRPSANTTYPRKALINEMGSGPCTEKKK	1078
Db	1018	MLEBPGSKTDLIDLADISLPCPQOEBEPLRLTPQNSDYTYPRKALINEMGSGPCTEKKP	1077
Qy	1079	SNSVPTSVHQAOPADIKYVAALGDSLITAVGARPNNSDDLPTSMGCLSMSTIGCDNLEBTH	1138
Db	1078	SNSVPTSVHKLQPADIKYVAALGDSLITAVGARARNSSDLMSWKGCLSMSTIGCDALBETH	1137
Qy	1139	TTLBNILKKNPYLLGSGTSTWEGTAGLNVAAEGARABDMPAQAMDLVERMKNSPDIULE	1198

Db	1138	TTLENNILKKFNFSIFGSTGTLETBETAGFNVAVEARAROMPAQARDLVERMKASTEINLE	1197
Qy	1199	KDKMLVTLFTIGVNDLCHYCENPEAHLATERYOHIOQALDLISEELPPAFVNVVE-VMEIA	1257
Db	1198	MDMKLTILFTIGSNDLCHYCNPEHSAEEVVOHIOALDILYEBLPAPFINVVDIMELA	1257
Qy	1258	SLYXGOGGKC-AMLAANNCTCLRHSQSSLEKEKELKVMNNLLOHGISFSFYMHQYRED	1316
Db	1258	GLHGOGGCHTALPPAOSTSCSLRHFPSSPYIGELKVYTNLQSDMSRLSYQEKYTRFD	1317
Qy	1317	FAVVVOEFPONTLTLPLNERGDTDLTFPSEDCPHFSDRGHAEMALAMNNMLPVPGRKTS	1376
Db	1318	FAVVVOEFPQNTLLPLDKGSTDPTFPSEDCPHFSERGHAEMALAMNNMLEPVGKHTTF	1377
Qy	1377	NNFPHSRAKLKCPSPESPYLYTLRNSKLLPDQAEAEVLYMAVPVAAG---VGLVYGI	1432
Db	1378	NNFNYNTRTKLCPSTESPPLYTLQNSLSPVQTEKASGAVPGLVSAAGGLLVGLIVGI	1437
Qy	1433	IGTVVMRCRRGRREDPPMSL	1453
Db	1438	LAVSLMSSFRRRQKSPPESTV	1458

Query Match	Best Local Similarity	Matches 1012;	Conservative 161;	Mismatches 263;	Indels 9;	Gaps 6;
Qy	5	PGIFL--LELLLLIGQTPQIHTSPKRSITLSEQLMPELTKNSPPCPNKLGVMNPSKV	68.0%; Score 5279.5;	DB 11;	Length 1450;	
Db	5	PGSVLVLLTLLLLIGQSPQIHSSGSGENTSQPOQVFFTLTKNPFPCPKKLEISVLSKV	70.0%; Pred. No. 0;			
Qy	63	HSLLKPSIKRVVAIIGNLEIRPPDGTG--DLEK--QDMTERQQQCMGMVTLSDIIRFSP	70.0%; Score 5279.5;	DB 11;	Length 1450;	
Db	63	HSLLKPSIKRVVAIIGNLEIRPPDGTG--DLEK--QDMTERQQQCMGMVTLSDIIRFSP	70.0%; Pred. No. 0;			
Qy	120	SVNPVCHTKRKYVPHDG--AEDLMIQAOELVRNMKENTLQDFQDWKLIVFSSNASQCY	70.0%; Score 5279.5;	DB 11;	Length 1450;	
Db	125	SVLMPTSPKSGKRGHTTIDEDIMIQAKELVRLKONPDLDFKDKMLITVLEFSNNSQCH	70.0%; Pred. No. 0;			
Qy	179	LCPSAQONGIAGGVDELGMGLVDYLQOEVRPAFVNIJDSLEVAVSQRQYHGTVLSPAPR	70.0%; Score 5279.5;	DB 11;	Length 1450;	
Db	185	LCSSDQQKRLHMGMENTSGVLVDLHREVPRAFVNIJDSLEVTMAQQHQETGFSPPAIR	70.0%; Pred. No. 0;			
Qy	239	CNCSSEETRLAKYVMQWSTYQEAANSLIASRRYQESFTYVQPPRYETTPSLHSDEPR	70.0%; Score 5279.5;	DB 11;	Length 1450;	
Db	239	CNCSSEETRLAKYVMQWSTYQEAANSLIASRRYQESFTYVQPPRYETTPSLHSDEPR	70.0%; Pred. No. 0;			

Db 245 CKCESEITKLSKAWOMWSYOAEWEDLLASSCFNKHETFAVVFOSFFSEV--ELPLERPS 302
 Qy 299 QDSTTLAMLMNMBEPAGEDEPLSYKHGRPMKCPSESPYLFSYNSNYTLRLKOPD 358
 Db 303 QDSTTLALRLWNSMMEVGRKDGTLNEAEKTKMKCSQESPYLFTYNSNYQAROLKEP 362
 Qy 359 KLEVRGAEIRCPDKPSDTPVTSVHRLKPADINVLGALDSLTAAGASTPGNVLDVL 418
 Db 363 KFMKGKTKFTCPDKDPSDSIPTVHRLRPADIKVIGAMGDSLTAAGASTPGNVLDVL 422
 Qy 419 TQYRGISWSVSGDENICTVTTLNIIAREPNSLKSFGVGTGKTSPPNAFLNQAAGRAE 478
 Db 423 TQYRGISWSVSGDETLETVTTLNIIAREPNSLKSFGVGTGKTSPPNAFLNQAAGRAE 482
 Qy 479 DLVQARLVDLMKNDRIHFOEDMKITLPIGNDLCPDNLVHSPDNFNDIKGAL 538
 Db 483 GLAAQAKLVSLMKDDKTTFQEDMKITLPIGNDLCSGNNLARSPTFTDNKTAL 542
 Qy 539 DILHAEPRAVNLVTVLEIVNLRELVOEKVYCPRMILSLCPVCLKPDSNSTELATLI 598
 Db 543 DILHAEPRAVNMVSTIELTRELNEPKVSCPRMILSLCPVCLKPDSNSTELATLI 602
 Qy 599 EPNKKTQEKTHQILIESGRYDTRDFVYVOPPEBNDMPKTSGLPDSNPFADPCFHS 658
 Db 603 ERNRQYOEETGKLIESGRYDTRDFVYVOPPEBNDMPKTSGLPDSNPFADPCFHS 662
 Qy 659 KSHSRASALMNMLEBVGOKTRHKEKNINTCPNOVOPPLRYKNSNGHGTMLPCR 718
 Db 663 KTHARSAIALMKMLEBVGKTRHKEKNINTCPNOVOPPLRYKNSNGHGTMLPCR 722
 Qy 719 DRAPSLHPTSVHLPADIQVVAALGDSLTAAGISGKRPDDPDVYQYRGISYAGD 778
 Db 723 EKAPSAISPTSVTLRPADIQVVAALGDSLTAAGISGKRPDDPDVYQYRGISYAGD 782
 Qy 779 GSENVTTPLNIIREFNRNLTVYAVGTGANDTNAFLNQAAGRAEMLSOVOTLMQX 838
 Db 783 KPELVNTPLNIIREFNRNLTVYAVGTGANDTNAFLNQAAGRAEMLSOVOTLMQX 842
 Qy 839 KDHARNFHEMDKVTYVLIGGSDLCVCTDONSILYSANFYDHLRNALDVLHREPVLYN 898
 Db 843 KNDTRNFHMDKVTYVLIGGSDLCVCTDONSILYSANFYDHLRNALDVLHREPVLYN 902
 Qy 899 LVDFLNPFTIRQVFLGNPDKCPVOQASVLCNVCYTLRENSOELARLEAFRAARSSREL 958
 Db 903 LVDFLNPFTIRQVFLGNPDKCPVOQASVLCNVCYTLRENSOELARLEAFRAARSSREL 962
 Qy 959 VSGGRYDTRDFSVLQPFQNTQLPVADGLPDTSEFADPCIHPOKFSHOLARLMTN 1018
 Db 963 VSGGRYDTRDFSVLQPFQNTQLPVADGLPDTSEFADPCIHPOKFSHOLARLMTN 1022
 Qy 1019 MLEPLGSKETLTLRAEMPTTCPTQNEPPLRTFRNSNYTPIKPAIENMGSDFLCTEWSA 1078
 Db 1023 MLEPLGSKETLTLRAEMPTTCPTQNEPPLRTFRNSNYTPIKPAIENMGSDFLCTEWSA 1082
 Qy 1079 SNEVPTSVHOLRPADIKVVAALGDSLTAAGARPNNSSDLPSTWRGLSMSIGDGNLETH 1138
 Db 1083 SNEVPTSVHOLRPADIKVVAALGDSLTAAGARPNNSSDLPSTWRGLSMSIGDGNLETH 1142
 Qy 1139 TTLPNILKKNPYLLGFTSTWEGTAGLVNVAEGARADMPAQAAMDIVERMKNSPDINLE 1198
 Db 1143 TTLPNILKKNPYLLGFTSTWEGTAGLVNVAEGARADMPAQAAMDIVERMKNSPDINLE 1202
 Qy 1199 KDMKLVTLFGVNDLCHYCENPRAHLATEYVQHIQOALDILSEELPRAFNVAEVELAS 1258
 Db 1203 KDMKLVTLFGVNDLCHYCENPRAHLATEYVQHIQOALDILSEELPRAFNVAEVELAS 1262
 Qy 1259 LVYOGGKGCAM-LAAQNNCTCLRHSOSLSEKOLKKNMNLQHGISFSFYWHQYOTREDP 1317
 Db 1263 LVYOGGKGCAM-LAAQNNCTCLRHSOSLSEKOLKKNMNLQHGISFSFYWHQYOTREDP 1322
 Qy 1318 AVVVOPEFNTLPLNERGDTDLTFPSSEDCFHSVDRGHAEALALMNMMLPEVGRKITSN 1377
 Db 1323 AVVVOPEFNTLPLNERGDTDLTFPSSEDCFHSVDRGHAEALALMNMMLPEVGRKITSN 1382

Qy 1378 NPTSPRAKLKCPSPESPYLTLNRSRLLPDOAEAEVLYMAVPAVAGVLYGIGTVV 1437
 Db 1383 NPTSPRAKLKCPSPESPYLTLNRSRLLPDOAEAEVLYMAVPAVAGVLYGIGTVV 1442
 Qy 1438 WRCRR 1442
 Db 1443 WRTVR 1447
 RESULT 3
 Q81UP7
 ID 081UP7 PRELIMINARY; PRT; 423 AA.
 AC 081UP7;
 DT 01-MAR-2003 (T-EMBLrel. 23, Created)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
 DE Similar to phospholipase B.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 NC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NX NCBL_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL, BC042674, AAH42674.1, -
 SQ SEQUENCE 423 AA; 47354 MW; A20A193491BA780C CRC64;
 Query Match 29.2%; Score 2270; DB 4; Length 423;
 Best local Similarity 99.8%; Pred. No. 176-160;
 Matches 422; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1036 MPITCTQNEPPLRTFRNSNYTPIKPAIENMGSDFLCTEWSKNSVPTSVHOLRPADIK 1095
 Db 1 MPITCTQNEPPLRTFRNSNYTPIKPAIENMGSDFLCTEWSKNSVPTSVHOLRPADIK 60
 Qy 1096 VVAALGDSLTTAVGARPNNSSDLPSTWRGLSMSIGDGNLETHTTPLNIIKKNPYLLGF 1155
 Db 61 VVAALGDSLTTAVGARPNNSSDLPSTWRGLSMSIGDGNLETHTTPLNIIKKNPYLLGF 120
 Qy 1156 STSTWEGTAGLVNVAEGARADMPAQAAMDIVERMKNSPDINLEKDKLVTLFGVNDLCH 1215
 Db 121 STSTWEGTAGLVNVAEGARADMPAQAAMDIVERMKNSPDINLEKDKLVTLFGVNDLCH 180
 Qy 1216 YCENPRAHLATEYVQHIQOALDILSEELPRAFNVAEVELASLYOGGKGCAMLAQNN 1275
 Db 181 YCENPRAHLATEYVQHIQOALDILSEELPRAFNVAEVELASLYOGGKGCAMLAQNN 240
 Qy 1276 CTCLRHSOSLSEKOLKKNMNLQHGISFSFYWHQYOTREDPAVVVOPFQNTLPLNER 1335
 Db 241 CTCLRHSOSLSEKOLKKNMNLQHGISFSFYWHQYOTREDPAVVVOPFQNTLPLNER 300
 Qy 1336 GBDTLTFPSSEDCFHSVDRGHAEALALMNMMLPEVGRKITSNPTSPRAKLKCPSPESPY 1395
 Db 301 GBDTLTFPSSEDCFHSVDRGHAEALALMNMMLPEVGRKITSNPTSPRAKLKCPSPESPY 360
 Qy 1396 LVTLRNSRLLPDOAEAEVLYMAVPAVAGVLYGIGTVVWRCRRGRREDPMSLRT 1455
 Db 361 LVTLRNSRLLPDOAEAEVLYMAVPAVAGVLYGIGTVVWRCRRGRREDPMSLRT 420
 Qy 1456 VAL 1458
 Db 421 VAL 423
 RESULT 4
 Q96DP9
 ID 096DP9 PRELIMINARY; PRT; 270 AA.
 AC 096DP9;
 DT 01-DEC-2001 (T-EMBLrel. 19, Created)
 DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)

DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
 DE Hypothetical protein FLJ30866.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Nishi T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H.,
 RA Sugiyama T., Irie K., Otsuki T., Sato H., Makamatsu A., Ishii S.,
 RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
 RA Kikuchi H., Kanda K., Matsuo K., Nakamura Y., Sekine M.,
 RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
 RA Sugano S., Nagahara K., Masuno Y., Nagai K., Isogai T.,
 RT "NEDO human cDNA sequencing project."
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK055428; BAB70920.1; -
 DR InterPro; IPR001087; Lipase_GDSL.
 DR Pfam; PF00657; Lipase_GDSL; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 270 AA; 30869 MW; AC53BC75FD09523 CRC64;
 Query Match 17.9%; Score 1392; DB 4; Length 270;
 Best Local Similarity 100.0%; Pred. No. 2, 2e-95;
 Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 491 MKNDRIHPOEDMKIITLFLFGNDLCPDNDLVHSPQNFNTDNGKALDILHAEVPRFV 550
 DB 1 MKNDRIHPOEDMKIITLFLFGNDLCPDNDLVHSPQNFNTDNGKALDILHAEVPRFV 60
 QY 551 NLTVLEIIVLRELQYKRYCPMILRSICPCVLKRDNDSTELATLIEFNKKRQETHQ 610
 DB 61 NLTVLEIIVLRELQYKRYCPMILRSICPCVLKRDNDSTELATLIEFNKKRQETHQ 120
 QY 611 LIISGRYDTEDEFTVWVQPFENVDMKTSGLPDNSFFAPDCFFHSSKSHRAALMN 670
 DB 121 LIISGRYDTEDEFTVWVQPFENVDMKTSGLPDNSFFAPDCFFHSSKSHRAALMN 180
 QY 671 NMLEPVGQKTRHKEFNKINITCPNOVQPLRTYKNSMOGHGWLPCRDRAPSALHPTSV 730
 DB 181 NMLEPVGQKTRHKEFNKINITCPNOVQPLRTYKNSMOGHGWLPCRDRAPSALHPTSV 240
 QY 731 HALRPADIOVVAALGDSL 749
 DB 241 HALRPADIOVVAALGDSL 259

RESULT 5

Q9D4Y6 PRELIMINARY; PRT; 394 AA.
 ID Q9D4Y6
 AC Q9D4Y6
 DT 01-JUN-2001 (Tremblrel. 17, Created)
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE 4930539A06RIK protein.
 GN 4930539A06RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Matsuda Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Okazaki H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Maehio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsch G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamaya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Monbaerts P.,
 RA Nordone P., Ring B., Ringwald M., Mazzarelli J., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitlaker C., Wilming L.,
 RA Wyszynski B., Yoshida K., Hasegawa Y., Kawai H., Kohenki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL; AK015993; BAB30072.1; -
 DR MGI; MGI:1922406; 4930539A06RIK.
 SQ SEQUENCE 394 AA; 44290 MW; E6D33EC17B6F0888 CRC64;

Query Match 14.7%; Score 1145; DB 11; Length 394;
 Best Local Similarity 59.8%; Pred. No. 1, 1e-76;
 Matches 234; Conservative 46; Mismatches 97; Indels 14; Gaps 7;

QY 1 MGRPGL--FLLELLLLGQGTQITSPRKSITLSEQ---LMPETLKNSPFCNPKUG 54
 DB 1 MELYPGVSPVGLLLLLGQGPQIHSGSGENTLAWQSQQVFW--TLKPFPECKEKLE 58
 QY 55 VNNPSKSVSLKPSDKIFVAALGNLEIPDPGTV--DLEK-QDTERPOOVCGWTVLS 111
 DB 59 LSVLSSEVSLKPSDKIFVAALGNLEIPDPGTV--DLEK-QDTERPOOVCGWTVLS 118
 QY 112 DIIRYSPSPVMPVCHGTRKRVIPHDGADLMIOAQLVNNKMNILQDFQDKLNVFF 171
 DB 119 DIIRHSPSPVMTGSPGKGTAVHTAEDLMIOAKELVRLKONPQDPEKDKLITVFF 178
 QY 172 SNASQCLCPSAQNLGAAGVDELNGVDLYLOEIPRAFNVLVDISEVAESRQHGTV 231
 DB 179 SNTSQCHLCPSAQNLGAAGVDELNGVDLYLOEIPRAFNVLVDISEVAESRQHGTV 238
 QY 232 LSPAPPCNCSERTFLAVVMQMSYQEAAMNSLLASRSSEBSFTVFPQPFYETPSL 291
 DB 239 FSPAPPCNCSERTFLAVVMQMSYQEAAMNSLLASRSSEBSFTVFPQPFYETPSL 297
 QY 292 HSEDPLQDSTTLAMVNNRMMEPAGEKDEPLSVKGRPMKCPQSPYLFYSRNSNYLT 351
 DB 298 KRSSP--QPTTLALINMSVMEPVGQGLNTATARKTKMCSSESPFLTYKNSNYQA 355
 QY 352 RLQKPDKLEVRGAEIRCPDPSPTVPS 382
 DB 356 RLKPYTKLQMKESSEFTCPDKNPSNIPPT 386

RESULT 6

Q9N2Z4 PRELIMINARY; PRT; 377 AA.
 ID Q9N2Z4
 AC Q9N2Z4
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
 DE Hypothetical 41.4 kDa protein.
 GN Y65B4B.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Br15col N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA None;
 RT "Genome sequence of the nematode C. elegans: a platform for
 investigating biology. The C. elegans Sequencing Consortium."
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.

```

RC STRAIN=Bristol N2;
RA Maupin R., Bauer C.;
RT "The sequence of C. elegans cosmid Y65B4B.R.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL/AC024847; AAF60857.1; -.
DR WormPep; Y65B4B.R.1; CE25539.
DR InterPro; IPR001087; Lipase_GDSL.
DR Pfam; PF00657; Lipase_GDSL; 1.
KW Hypochemical protein.
SQ SEQUENCE 377 AA; 41400 MW; 2A93D3842C80E6A5 CRC64;

Query Match
Best Local Similarity 8.2%; Score 636; DB 5; Length 377;
Matches 132; Conservative 71; Mismatches 122; Indels 16; Gaps 8;

QY 376 SDTVPTSVRLKRLADINVLGALGDSLTAGAGSTGPNVLDVLTQYRGLSMVSGDENG 435
DB 37 SKNTPTSASVSRPTDIKVLGALDLSLTAAAGAPKGDPLAVILQYRGLAFQCGGSHSLD 96
QY 436 TTTTLNILREFNPSLKGFSVGTGKETS-PNAF---LNOAVAGRAEDLPVQARRLVDM 491
DB 97 EHTVAVNLKKEPSNLMGISTGIG--SANYWEVSKLNQVPGAEALDITQARALVQII 153
QY 492 KNDTRIHFQEDWKITLFTIGNDLCPDNDLVH--VSPQNTNIGKALDILAEVPRAF 549
DB 154 QSHKEIDYKTDKTLINVFICANDMCAYCNDRENGPSKALWKQNVITAIQILKDNLPRTI 213
QY 550 VNLVTVLEIVNRELVOEKVYCPRMILKSLCPVLKFDNSTELATLIEFNKKFOEKT 609
DB 214 VSMTGMFDAMMLQI-DHDKYFCDGLHVE-CPCEKKNKDPNDTISAACHL---YMDAQ 268
QY 610 QLESGRYDREDPTVVVQPFNVDPKTS-BGLPNSFPAEDCPHFSSKSHRASAL 668
DB 269 EIDDSGIFDSTDTFVQVPPFNGITVPLKPDSEVNLDFAPDCFFSKLGHANVAKIL 328
QY 669 MNMLEPVGOKTRHKEFN-KINITCPNOVQPLRTYKNSM 708
DB 329 MNNIQPVGSKNQVNLSDPTILNCPDTCFPPRTTKNSV 369

RESULT 7
Q8K255 PRELIMINARY; PRT; 186 AA.
AC 08K255;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-OCT-2002 (Tremblrel. 22, Last annotation update)
DE Hypochemical protein.
DE Moe musculus (Moose).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL/BC033606; AAH33606.1; -.
KW Hypochemical protein.
SQ SEQUENCE 186 AA; 21471 MW; 56153A626A22315 CRC64;

Query Match
Best Local Similarity 8.2%; Score 635; DB 11; Length 186;
Matches 116; Conservative 18; Mismatches 26; Indels 0; Gaps 0;

QY 1289 QELKKVNNLQHGISFSYHQYTOREDFAVVVQPFQNTLPLNERGDTLTFPSDCP 1348
DB 2 QELKKLNNVQSGISLSYHRYMEREDFAVTVQPFPRNFTPLNERGDLTFPSDCP 61

```

```

QY 1349 HFSDRGHAEMALAMNNMLPVGKRTSNFTSHRAKLCPSPESEPLYTLNRSRLLPQ 1408
DB 62 YFSDRGHAEMALAMNNMLPVGKRTSNFTSHRAKLCPSPESEPLYTLNRSRLLPQ 121
QY 1409 AEEPAEVLVAVPVAAVGLVVGIIIGTVVRCRGRRED 1448
DB 122 AEEPANLVAVPVAAVGLVVGIIIGTVVRCRGRRED 161

RESULT 8
Q23119 PRELIMINARY; PRT; 382 AA.
AC 023119;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE W02B12.1 protein.
DE W02B12.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderiinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Swinburne J., Almscough R.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Gene sequence of the nematode C. elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z66521; CAA91393.2; -.
DR WormPep; W02B12.1; CE03761.
DR InterPro; IPR001087; Lipase_GDSL.
SQ SEQUENCE 382 AA; 41772 MW; 6430145B8525CA3 CRC64;

Query Match
Best Local Similarity 8.1%; Score 628.5; DB 5; Length 382;
Matches 139; Conservative 57; Mismatches 119; Indels 37; Gaps 8;

QY 376 SDTVPTSVRLKRLADINVLGALGDSLTAGAGSTGPNVLDVLTQYRGLSMVSGDENG 435
DB 35 SKNTPTSASVSRPTDIKVLGALDLSLTAAAGAPKGDPLAVILQYRGLAFQCGGSHSLD 94
QY 436 TTTTLNILREFNPSLKGFSVGTGKETS--PNAFLNOAVAGRAEDLPVQARRLVDMKND 494
DB 95 EHTVAVNLKKEPSNLMGISTGIG--SANYWEVSKLNQVPGAEALDITQARALVQII 154
QY 495 TRIHQEDWKITLFTIGNDLCPDNDLVH--VSPQNTNIGKALDILAEVPRAF 549
DB 154 QSHKEIDYKTDKTLINVFICANDMCAYCNDRENGPSKALWKQNVITAIQILKDNLPRTI 213
QY 550 VNLVTVLEIVNRELVOEKVYCPRMILKSLCPVLKFDNSTELATLIEFNKKFOEKT 609
DB 214 VSMTGMFDAMMLQI-DHDKYFCDGLHVE-CPCEKKNKDPNDTISAACHL---YMDAQ 268
QY 610 QLESGRYDREDPTVVVQPFNVDPKTS-BGLPNSFPAEDCPHFSSKSHRASAL 668
DB 269 EIDDSGIFDSTDTFVQVPPFNGITVPLKPDSEVNLDFAPDCFFSKLGHANVAKIL 328
QY 669 MNMLEPVGOKTRHKEFN-KINITCPNOVQPLRTYKNSM 708
DB 329 MNNIQPVGSKNQVNLSDPTILNCPDTCFPPRTTKNSV 369

RESULT 9
Q01300 PRELIMINARY; PRT; 374 AA.
AC 001300;
DT 01-JUL-1997 (Tremblrel. 04, Created)

```


DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE F09C8.1 protein.
 GN F09C8.1.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA McMurtry A.A.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "genome sequence of the nematode C.elegans: A platform for
 RT investigating biology."
 RT Science 282:2012-2018(1998).
 DR EMBL; Z68132; CA922221.2; -.
 DR WormPep; F09C8.1; CE31470.
 DR InterPro; IPR001087; Lipase_GDSL.
 DR Pfam; PF00657; Lipase_GDSL.
 DR PROSITE; PS50241; LIPASE_GDSL.
 DR PROSITE; PS50241; LIPASE_GDSL.
 SQ SEQUENCE 374 AA; 41921 MW; 553E0DB7EE3A41EC CRC64;
 Query Match 8.1%; Score 625.5; DB 5; Length 374;
 Best Local Similarity 39.8%; Pred. No. 5.3e-38;
 Matches 139; Conservative 60; Mismatches 111; Indels 39; Gaps 9;
 QY 376 SDTVPSTVHRLKPADINIVGALGDSLTPAGAGSTPGNVLDVLTQYRGLSMVSGDENIG 435
 DB 37 SKVPTNVNSVRPADITLALGDSLTPAGAGSE--DPVAVVQYHGLAFQAGDITLE 94
 QY 436 TTTTLNILEFNPGLKGFVSGTGTSPN---AFINQAVAGRAEDLPQVQARLVDM 491
 DB 95 EHTVTPILTKKYNDPVGYSNGIG---SPVWEIARLVNMPGANAKDLFGQARQLVQL 151
 QY 492 KNDTR-IHPQEDMKIITLFTGNDLCPQNDLVHYSQNTDNGKALDILHAEPVAPV 550
 DB 152 QQHTEVVNMKEEDWLTINIFIGNDICGYCRKPVDSYNGCAODIKQAVQIITVNVPRVIV 211
 QY 551 NLVTLEIVNIRELYQEKVYCPMILRSLCPCLKEDDNSTELATLIEFKKRFQ 606
 DB 212 SLTGMLHEMLRQY-DIGHMFCQR-LHNDGCGC-----ESKAKFTDIDR 254
 QY 607 -----KTHQLIESGRYDTRDEFTVVVQPFENVDMPTSEGLPDNSFPAPDCFHFSK 659
 DB 255 QACDYVNYKEQJLETDGTFEKNDFTVYVQPMFQDTLIPRMENGRKPTQKFPAPDCFHFSQW 314
 QY 660 SHSRASALNMNMLEPVGQKTRHKEFNKI-NITCPNQVQPFRTYKNS 707
 DB 315 GHALVSTYLMNNTLQPVGSKSYVSNMSVPLQTLACPDACPFRITPNS 363
 RESULT 10
 Q8MX08 PRELIMINARY; PRT; 398 AA.
 AC Q8MX08;
 DT 01-OCT-2002 (TReMBLrel. 22, Created)
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Hypothetical protein Y73B6BL.37.
 GN Y73B6BL.37.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Briscol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA Waterston R.;

RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium."
 RT Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Briscol N2;
 RA Graves T.;
 RT "The sequence of C. elegans cosmid Y73B6BL."
 RT Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Briscol N2;
 RA Waterston R.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC084197; AAM44393.1; -.
 DR WormPep; Y73B6BL.37; CE29931.
 DR InterPro; IPR001087; Lipase_GDSL.
 DR Hypothetical protein.
 SQ SEQUENCE 398 AA; 44003 MW; E6027C51E47E2B76 CRC64;
 Query Match 7.9%; Score 616; DB 5; Length 398;
 Best Local Similarity 36.0%; Pred. No. 3e-37;
 Matches 133; Conservative 74; Mismatches 120; Indels 42; Gaps 10;
 QY 374 DP-----SDTVPSTVHRLKPADINIVGALGDSLTPAGAGSTPGNVLDVLTQYRGLSMV 428
 DB 30 DPLLYKKSQGVPTSAISVSRPDIKITIGALGDSLTPAGAGAPGDDPLAVILQYRGLAFQC 89
 QY 429 GGDENIGTVT-----LANILREFNPGLKGFVSGTGTSPNAP 467
 DB 90 GGDSDIDENHTVASENFTNLQTSAPAIMQLADVKKFSPVIMGYSIGIG---SANVW 146
 QY 468 ----LNOAVAGRAEDLPQVQARLVDMKNDRIHPOEDMKIITLFTGNDLCPQNDLV 523
 DB 147 EYSKLQAVPAGAAADITQARLVQIMQNHDDIDKNMKLVNFIQANDMCRYEKE 206
 QY 524 H--YSPQNFQDNGKALDILHAEPVAPVNLVTLEIVNIRELYQEKVYCPMILRSLC 581
 DB 207 NGHISAKMKQVNTTAIQLKDLPRITVSMGMPFMRMLRQ-DHDKFQDGLHAF-C 264
 QY 582 PCVLKEDDNSTELATLIEFKKRFQEKTHQLIESGRYDTRDEFTVVVQPFENVDMPTSE 640
 DB 265 PCSKKSFNTNDVDSQACHL---YMDAQOEIQDSGIDFTDDBDFVLPQPFNGITIPPLK 321
 QY 641 EGLPDNSFPAPDCFHFSKSHSRASALNMNMLEPVGQK-TRHKEFNKINITCPNQVQ 699
 DB 322 DGEVNLDMFAPDCFHFSRLGHANVAGHNMNIVQPVSKKHIVLSDPTIPLNCPATCP 381
 QY 700 FLRTYKNSM 708
 DB 382 FIRTITNSV 390
 RESULT 11
 Q9VW80 PRELIMINARY; PRT; 424 AA.
 AC Q9VW80;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE CG7365 protein (GH22081p).
 GN CG7365.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.C., Chame M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Avril J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Baller R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Maye A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fosløe C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
 RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Kapten G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mleishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler P., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*,"
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Chame M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Garxin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuno J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.,
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AEO0314; AAF9069.1; -
 DR EMBL: AY069174; AAL3919.1; -
 DR FLYBase: FBgn0036939; CG7365.
 DR InterPro: IPR001087; Lipase_GDSL.
 DR Pfam: PF00657; Lipase_GDSL.
 DR PROSITE: PS01098; LIPASE_GDSL.
 SQ SEQUENCE 424 AA; 48479 MW; 15DC6229188E171 CRC64;

Query Match 7.5%; Score 584; DB 5; Length 424;
 Best Local Similarity 38.5%; Pred. No. 8.2e-35;
 Matches 117; Conservative 57; Mismatches 116; Indels 46; Gaps 12;

QY 369 RCPDPSDPTVPSVRLKPADINVTGALGDSLTAGAGSTPCGVNLDVLTQYRGLSWSV 428
 DB 88 RSPER-----PTSVRLRPGDIDIVGAMGDSLTAGAGIPAT--NLHTVVENRGVWMI 139
 QY 429 GGDENIGTVTLTANILREPNPSLKGPSVGTGKTSPPAFINQAVAGRADLVQARRLV 488
 DB 140 GGQYDMRKVTLTNPILKEFPNLYGYATKQISTDRSRDVAELAMSKDPMKVKLV 199
 QY 489 DLAKNDTRIHFDQWKITLFIGNDLCPDNDLVHY-SPQNTD-----NIGKALDILHA 543
 DB 200 RRMQRDPVPMVMSDQKLVTLFIGNN---DCTDIDCYPEPEKVDHMERMLKTYRYLRD 256
 QY 544 EVPRAPVNLTVLEIVNLRRL-----YQKKVYCPRMILRSICPVLFKFDNNGSTE 593
 DB 257 NVPRMLNLTIVPA---PILRFLTNLTGILPICYG-----TLRFECPLM--GKQKGQ 302

QY 594 LATLIEFNKKFOEKTHOLIESGRYDREFTVVVQPFENVMPKTSSEGLPNSFPAPDC 653
 DB 303 LDVLEGIMKMKIKQDEIANRBEFNT--EPFTIVQPSQGPDPFRRSQGTDIRPSEDC 361
 QY 654 FHSSKSHSRASALNNMLPEVQKGT--TRKFEKNKINTCNOVOPPLRTYKNS 707
 DB 362 FHLSQGHAAANNSIWNMLLEPGHKSQGFATHLE---TFRCSEMRPFILITRENS 414

RESULT 12

ID Q9VMM8 PRELIMINARY; PRT; 447 AA.
 AC Q9VMM8
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE CG11029 protein.
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; *Drosophila*.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams W.D., Celniker S.E., Holt R.A., Evans C.A., Gccayne J.D.,
 RA Anantides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.C., Chame M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Avril J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
 RA Baller R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Maye A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.B., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fosløe C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
 RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Kapten G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mleishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler P., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*,"
 RL Science 287:2185-2195(2000).
 DR EMBL: AEO03611; AAF52285.1; -
 DR FLYBase: FBgn0031735; CG11029.
 DR InterPro: IPR001087; Lipase_GDSL.
 DR Pfam: PF00657; Lipase_GDSL.
 DR PROSITE: PS01098; LIPASE_GDSL.
 SQ SEQUENCE 447 AA; 51261 MW; BBD3F53DCB35DEB CRC64;

Query Match

6.6%; Score 515; DB 5; Length 447;

Best Local Similarity 31.1%; Pred. No. 1.3e-29;
Matches 137; Conservative 73; Mismatches 155; Indels 76; Gaps 14;

QY 293 SEDPRLODS-----TTLAHMLNMRMEAGKDEPLSYKGRPM 331
Db 46 SYDRVRENGIQOYTIDIDQLRLFLNTRQTLTKMALNN-----IEALSRRGRREG 96
QY 332 KCPQSESPYLFSYNSNVLTRLOKPODKLEVBEGAIKPCPKDP-SDTVPTSVRLKPAD 390
Db 97 K-----LQAP-----VSKVPELCPNNTRRSFSPSTISHLRPGD 131
QY 391 INVIGALGDSITAGNGAGSTPGVNLVLTQYRGLSWSVGSDENIGVTTLANILREPNPS 450
Db 132 IDIIAAGDSISAGNGILS--NNAI DMINEFRGLTFSGGGLANRRFVTLNLIKIFNPK 189
QY 451 LKGSVGTGKETS-PNAPLNOVAGGAAEDLPVQARLVLMKRDTHIQEDMKITLTF 509
Db 190 LYGFVAVNSLVINRRSLRNLIAEPMISRDLPQARVILDLRRDRVDMKRWKLLTVY 249
QY 510 IGGNDLDFGNDLVHY-SPQNFDT---NIGKALDILHAEPRAVNLVTVLEIVNREL 564
Db 250 VGANDI---GSDLCMTTPGSLDQHARDLRQAFRLRDHVPRLINLIVPNIPVLST 306
QY 565 YQEKRYVCPMILASLCPVLFKPDNSTELATLIEFNKKQOKTHQIIESGRYDREDPT 624
Db 307 MTKVPLQC-FVVRHVGHCHLINDRLNTEFNERMDTLTRMQDLMEIARLPEF-RQDFA 364
QY 625 VVVOFFENVDMPTSGSLPDNSFFAPDCFFFSKSHSRASALMNNMLEVQCK---T 680
Db 365 IVAHPMLTKLTAPLPDGSTDWRFPSHDCFFHSORGHAIISNLMNSMLLPDDQKPRPSV 424
QY 681 TRHKFNKINTCNVOVQFL 701
Db 425 VPELFE---RVVCPTAEQPIYL 442

RESULT 13

Q950L1 PRELIMINARY; PRT; 460 AA.

AC Q950L1; (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Hypothetical 51.9 kDa protein.
GN P13H8.11.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea;
OC Rhabditiida; Pelodertinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Briscot N2;
RX MEDLINE=99069613; Pubmed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Briscot N2;
RA Ding H.;
RT "The sequence of C. elegans cosmid P13H8.";
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Briscot N2;
RA Waterston R.;
RT "Direct Submision.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U23137; AA02483.1; -
DR WormPep; P13H8.11; CE29264.
DR InterPro; IPR001087; Lipase_GDSL.
DR Pfam; PF00657; Lipase_GDSL.1.

KM Hypothetical protein.
SQ SEQUENCE 460 AA; 51861 MW; 5D4108B1657319E1 CRC64;

Query Match 6.0%; Score 469.5; DB 5; Length 460;
Best Local Similarity 29.6%; Pred. No. 3.2e-26;
Matches 120; Conservative 72; Mismatches 174; Indels 39; Gaps 14;

QY 358 DKLEVBEG-AEIKCPDKDPEDTVPTSVHRLKPADINVIGALGDSITAGNGAGSTPGVNL 416
Db 43 DSEIEYPMGFNPFCP--QPIKSSSEVHQHPQIGVIALGDSVSAQAAS--SILD 98
QY 417 VLTQYRGLSWSVGSDENIGVTTLANILREFNSLGFVGTGKETSPPNAPLNOVAGGR 476
Db 99 LFDQFCVSVFTGDDVYLNQATFINIPRFAPRIKQ---GSSDFQKFDYFMALPGSF 155
QY 477 AEDLPQARLVLMKRDTHIQEDMKITLIEGNDLDFGNDLVHYSPQNFDTNIGK 536
Db 156 SSELPRQALKLASTLKKLGRETSDTWKFVNIPIGNDLGNINNETTFGPEFFGSLHS 215
QY 537 ALDILHAEPRAVNLVTVLEIVNRELVOEKKV--YCPMILRSCLPVLFKPDNSTEL 594
Db 216 ALTIITQTNVKKVFN--IMPIVVKIHQAHLKSF-C-EFSRKTCSCTIFELNEKEYQ- 270
QY 595 ATLIEFNKKQOKTHQIIE--SGRYDREDPTVVVOFFENVDMPTSGSLPDNSFFAPD 652
Db 271 ---NIGKQDEQDNEVEQFNQKGNSSFFAVIAPADMLKSI-P-LKNQPNIGLLAD 325
QY 653 CFFFSKSHSRASALMNNMLEVQCKTRHKFNKINTCNVOVQFL 707
Db 326 CFHLSPLAHIDIAKQIWKGLFEPIDQKT---ITNLSVGFDRFVCPPEVCEPYLRITONS 381
QY 708 MGHGTWLPQRD---RAPSLHPTSVHRLKPADIOVVALGDSL 748
Db 382 ENCE---PSREKRFILVPSAFASSPGLSMPTMIFVFLFGIAL 422

RESULT 14

Q21799 PRELIMINARY; PRT; 348 AA.

AC Q21799; (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE R07B7.8 protein.
GN R07B7.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea;
OC Rhabditiida; Pelodertinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Harris B.R.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=99069613; Pubmed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z75955; CAB00118.1; -
DR WormPep; R07B7.8; CE06270.
DR InterPro; IPR001525; C5_DNA_meth.
DR Pfam; PF00657; Lipase_GDSL.
DR PROSITE; PS00095; C5_MTASE_2; 1.
SQ SEQUENCE 348 AA; 39489 MW; 91F52DC5139363E0 CRC64;

Query Match 5.2%; Score 406.5; DB 5; Length 348;
Best Local Similarity 32.2%; Pred. No. 1e-21;
Matches 117; Conservative 66; Mismatches 133; Indels 47; Gaps 11;
QY 1060 IKPAIENWGSDFCTEMKASNSV-----PTSHQARLPADIKVVALGDSLTAVGARPV 1113

Db 15 VKKALE-----KYLNSIDPHEIYDPDVMMIKPHIRIVIGAMGDSLT--IGSRAB 61
QY 1114 NSSDLTPTSMWGLSWSIGGDNLEHTTLPLNLIK-----KFNPLYLGSTSTWEG-TAGLNV 1168
Db 62 NIVGQORIPFGNPFPTGMDPEVDRLVYVNI FRIIAEKTGNKLPFGSGTGIDYBENTGLNV 121
QY 1169 AAGCARADMPAQAAMDIVERMKNSPDINLEKMWLVTLFYGVND---LCHYCENPEAHLA 1225
Db 122 AIGGMSDDILRPAKEIVSRITANKKEINENDMKLSLWIGTNDVGLGRLEDPFP-V 179
QY 1226 TEYVOHIQALDILSEELPRAFYVNVVEWELASLYQGQGKCMALAAQNNCTCLRHSQS 1285
Db 180 DEYKSHIEKGLLYKENLPRTIVISVGMFPAQLQEAQ-----SILKNGKRA 226
QY 1286 LEKQELKVVNMVNIQHGSFSFYHQYTOR---EDFANVYQPF-FQNTLTPLNERGDTLT 1341
Db 227 RIVENQKLD-DLSDGGRNVSYDFQNNHNSNDFYVQPFATEYTDSYRDEHGKYNPT 285
QY 1342 FFESEDFHFESEDRGHAEMALAMNNMLEPVGRKTTNNFTSRAKLCSPSPSYLYTLRN 1401
Db 286 FYASDLFHLISKFGHVAIAKHYMLNLPVGEKTKRADLGDTKIYELNKNCLITVGN 345
QY 1402 SRL 1404
Db 346 SKM 348

RESULT 15

P90862 PRELIMINARY; PRT; 528 AA.
AC P90862;
DT 01-MAY-1997 (T-EMBLrel. 03, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (T-EMBLrel. 23, Last annotation update)
DE Hypothetical protein F36A2.9a.
GN F36A2.9a.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Lemard N.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology";
RL Science 282:2012-2018(1998).
DR EMBL; Z61077; CAB03072.2; -.
KW Hypothetical protein.
SQ SEQUENCE 528 AA; 59421 MW; 1B34202465E16261 CRC64;

Query Match 4.6%; Score 360.5; DB 5; Length 528;
Best Local Similarity 29.9%; Pred. No. 5.5e-18;

Matches 103; Conservative 59; Mismatches 134; Indels 49; Gaps 13;

QY 370 CPDKDSDVPTVTHRLKPADINVIGALGDSLTAAGNAGSTPGNVLDVLTQYRGLSMWVG 429
Db 130 CPRIKTELTGIVNGINSPEDITIIAMGDALTGIGLWPN---ADI--EFGASFPFG 183
QY 430 GDEINIGTITLANILREFNLSKFSVGTGKTSNPAFLNQAVAGRAEDLPVQARRLYD 489
Db 184 GDSITIGLITIPILREBFSKLVGVSHGAGDL-PNHQLNVAVTGATTEDELPQARRLTR 242
QY 490 LMKNDRIHQEDWKKITTLFISGNDLCPNDLVHVSPOFTNIGKALDILHAEVPRAF 549
Db 243 RLKKELELDVHNEMIMIIITIGTEELCSRCGPSY-----DNIRKAIHLOIEIPKA- 294
QY 550 VNLVTLEIYNLELVQEKVYCPRMILRLSCVLRKPDNSTELATLIEFNKKFQE--- 606

Db 295 --LVVLLGPVHSSFHEQKS-----NLLKARCAC-----SRDQTE-GFMYDVSRKWSKVMR 342
QY 607 KTHQLIESGRYDREDPTVYVQPFENVDMKP-TSEGLPDNSSFADCFHFSKSHSRAA 665
Db 343 DVQKFEVNG-----VTARPTFMISYPMVLTITSRYPSSGLFIRDKPLNRRGHNYAT 393
QY 666 SALMNM-----LEPVGQKTRHKEFNKINITCPNOVOPFLRTYKN 706
Db 394 KMLNRLIGDLYNLSSATL-----SQDNYFCPSVCGCPFRITYEN 433

Search completed: January 6, 2004, 18:59:17
Job time : 108 secs

THIS PAGE BLANK (USPTO)